Vol. 12 JUNE, 1941 No. 2

CHILD DEVELOPMENT

PUBLISHED QUARTERLY BY THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT
NATIONAL RESEARCH COUNCIL
2101 CONSTITUTION AVENUE
WASHINGTON, D. C.

Made in United States of America

CHILD DEVELOPMENT

Vol. 12

JUNE, 1941

No. 2

PUBLISHED QUARTERLY BY THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT
NATIONAL RESEARCH COUNCIL
2101 CONSTITUTION AVENUE
WASHINGTON, D. C.
Made in United States of America

SOME SEX DIFFERENCES OBSERVED IN A GROUP OF NURSERY SCHOOL CHILDREN

JEANETTE B. McCAY AND MARIE B. FOWLER[1]

In a study of the health and development of 31 girls and 35 boys, enrolled in the junior nursery school of the New York State College of Home Economics at Cornell University, during the five years, 1932-37, some interesting sex differences were noted.

Although these differences were too small to be statistically significant ($D/\sigma d$ was not equal to 3), their direction is of interest as giving further evidence in line with the results of others' studies. An increasing number of reports in the fields of biology, psychology, physiology, even nutrition point toward constitutional variations between the sexes and when brought together in summaries such as that of Johnson and Terman (7), present a rather consistent view.

Norms of the behavior which was observed including sleep, elimination, eating and nervous behavior at the noon-meal in school, as well as of growth, of colds and other sicknesses have been made (11).

Records were kept during the time which the children were enrolled in the nursery school. This totaled over 13,000 days for the 66 children, averaging about 200 days during the school year for each child. Infant histories and physical examinations were also analysed.

The median age of the boys and girls when school began was the same, 2 years 7 months, with the middle 50 per cent of each sex ranging between 2 years 4 months, and 2 years 10 months. Physical examinations showed the boys and girls equally well developed and equally healthy, with the boys as a group taller and heavier than the girls.

Eating Behavior

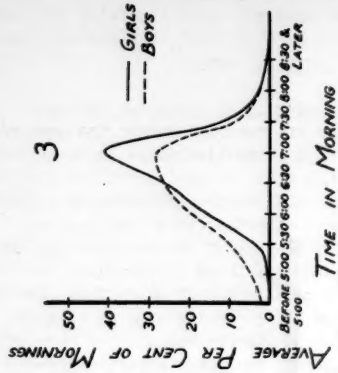
At the noon-meal in the nursery school, three measures of eating behavior were used: the time in minutes required for finishing a meal; the amounts of food eaten (measured in arbitrary units); and the "efficiency" in eating - a score representing the relationship between the amount eaten and the time. The use of these measures has already been described (12).

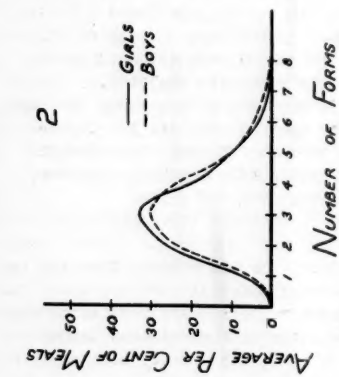
While the boys took practically the same amount of time for their meals as the girls, they tended to eat more food which gave them higher efficiency scores. Curve 1 in the chart compares the distributions for the girls and boys.These percentage distributions for each group were gained by averaging the percentages of the individual distributions at similar points.

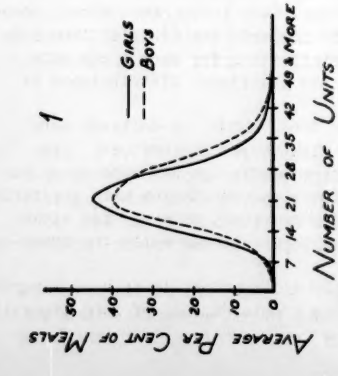
The selection of second servings of food, which the children were allowed to make on their own volition, also showed differences. The boys helped themselves to more of the breadstuffs and desserts than the girls although the amounts of protein foods and vegetables were similar. One might infer that the requirements for proteins, minerals and vitamins were alike but that the boys needed more calories which the breadstuffs and desserts supplied.

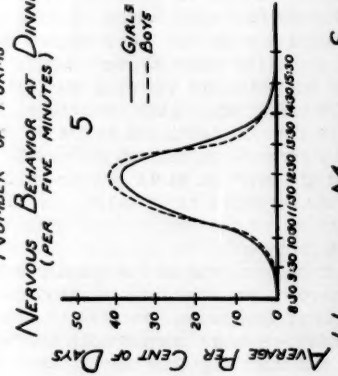
Since the boys as a group were taller and heavier, it appears natural that they would eat more at a meal. This greater amount of food probably

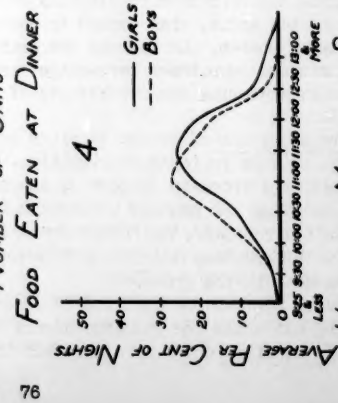
[1]*From Department of Family Life, New York State College of Home Economics at Cornell University, Ithaca, New York.*











Differences in Behavior Shown by Boys and Girls

Hours and Minutes of Sleep in Twenty-Four Hours

is also related to the higher and more constant basal metabolic rate and higher and more constant body temperatures which have been observed among boys (13) (2) (3).

Nervous Behavior

During four years, nervous behavior was recorded during the noon meal. The term "nervous behavior", as used in this study, refers to those restless movements, often involuntary, which serve no useful purpose and to which individuals often resort at times of strain or uncertainty. Some of the movements may be habitual, or they may appear rarely.

The method developed by Olson (14) was used. This consisted of recording each type of nervous behavior which a child was observed to make during an interval of five minutes. The forms used by Olson were revised and extended to cover the kinds of behavior observed in these children. Sixteen forms were studied, the names referring to the parts of the body involved in the movements (10). For the most part, observations covered the entire meal. At the end of the meal, the median number of forms per five minutes was listed for each child and the proportion of time during which different forms were observed.

Curve 2 shows that the median number of forms of nervous behavior per five minutes ranged from one to eight forms. Distribution curves are given for the girls and boys and indicate somewhat more restless movements by the boys. Although this difference for five minutes does not seem great, a full meal time would increase it considerably. And if this same difference is characteristic in other activities, then a longer space of time such as a day or week might yield very large differences, indeed.

These results are contrary to those of Olson's study where the girls in the first six grades of school gave higher averages in nervous habits than the boys. However, Olson did not use the large body, feet and leg movements which were included in these observations.

Just as the larger intake of food by the boys could be related to sex differences in size, in metabolism and body temperatures, more restless movements may also be related to the higher muscular tension at rest shown by boys and the greater physical activity of boys.

There was little difference between the boys and girls as to the various forms of behavior which were most habitual. The boys showed somewhat more hand to ear movements, vocalizing and repetition of movements made by other children. In certain years, the girls tended toward more hand to hair movements.

Sleeping Behavior

Records were kept of the children's sleep at home and at school. Of course it is difficult to tell accurately whether another person is awake or asleep, so the intervals for recording the time were made rather wide: 15 minutes for naps, half-hour intervals for sleep at night and hourly intervals for sleep during the 24 hours. Even so the trends indicate sex differences in the group.

Curve 3 gives the distributions for the times of getting up in the morning and shows the tendency of the boys to get up earlier than the girls. No differences were apparent in sleep at nap, nor in the bed-

time nor the time required for falling asleep at night.

As would be expected from the earlier rising hour, the boys tended to have shorter hours of sleep at night than the girls which is indicated in Curve 4.

Shorter night sleep also meant less sleep in 24 hours for the boys than for the girls, though the differences shown in Curve 5 are not so large.

There are few references to such sex differences in sleep, although Giddings (6) writes, "We have observed that, as a sex, females are sounder sleepers than males, not only do they get to sleep more quickly, but sleep more quietly throughout the night." And Erwin (5) concludes from her studies of 409 children that "slight sex differences reveal that girls sleep longer than boys."

Other Behavior

Differences were not so evident in the other types of behavior studied, although in elimination there was a slight indication of more bed-wetting and more times taken up for toilet at night for the boys. In addition, the boys showed a slightly higher percentage of time with two and more bowel movements a day than the girls and the boys' movements came more often in the morning than did the girls.

In rate of growth or amount of sickness, there were no differences apparent between the girls and boys of this group.

DISCUSSION

These differences of a higher food intake, more restless movements and an earlier rising hour, resulting in less sleep at night and during 24 hours for the boys are in line with many of the sex differences which have been reported. In addition to greater size, there are the tendencies of the male to be more aggressive and to show more extravert behavior than does the female. Differences in glands and internal secretions, we have learned may account for many differences in behavior (1). For instance, Bize and Moricard as reviewed by Marquis (8) noted that injections of the male hormone testosterone gave marked changes in the personality of young boys with hypogenitalism or adiposity. "There was an impression of increased strength and intellectual curiosity, and an increased aggressiveness, shown by a detachment from the family group and a greater tendency to assertiveness."

We have mentioned the tendency of the male to have a more constant and higher body temperature; to have a higher basal metabolic rate and higher muscular tensions while at rest. It seems fair to take the view that while there are large individual variations, males tend to be larger, more active, more tense, more dominant, perhaps "higher geared" organisms than are females. Perhaps it is this "higher geared" effect which accounts for the greater mortality of males in later life, a fact which is true of both the rat and man (4) (9).

SUMMARY

Sex differences observed in a group of children, 31 girls and 35 boys, enrolled during the years 1932-37 in the nursery school of the New York State College of Home Economics at Cornell have been reported. These differences were small, but in general the males were distinguished from the females by taller, heavier bodies, more food eaten at dinner, particularly the calory foods of breadstuffs and desserts, and more restless movements during the noon-meal. The boys of the group tended to awaken earlier in the morning than did the girls thus shortening their night's sleep and lowering the total amount of sleep gained during 24 hours.

REFERENCES

(1) Allen, Edgar, (ed.): Sex and internal secretions. Baltimore, Williams and Wilkins Co., 1939, pp. xxxv - 1346.

(2) Bayley, Nancy and Stolz, Herbert R.: Maturational changes in rectal temperatures of 61 infants from 1 to 36 months. Child Develop., 1937, 8, 195-206.

(3) Boothby, Walter M. and Paulson, Donald L.: Energy metabolism. In Luck, J. M., and Hall, V. E. (eds.) Ann. Rev. Physiol., 1940, 169-180.

(4) Ciocco, Antonio: Sex differences in morbidity and mortality. Quart. Rev. Biol., 1940, 15, 59-73 and 192-210.

(5) Erwin, Doris: An analytical study of children's sleep. J. Genet. Psychol., 1934, 45, 199-226.

(6) Giddings, G.: Child's sleep—effect of certain foods and beverages on sleep motility. Am. J. Pub. Health, 1934, 24, 609-614.

(7) Johnson, Winifred B. and Terman, Lewis M.: Some highlights in the literature of psychological sex differences/published since 1920. J. Psychol., 1940, 9, 327-336.

(8) Marquis, Donald G.: Physiological psychology. In Luck, J. M., and Hall, V. E. (eds.) Ann. Rev. Physiol., 1940, 433-461.

(9) McCay, C. M. and Crowell, Mary F.: Prolonging the life span. Sci. Month., 1934, 39, 405-414.

(10) McCay, Jeanette B.: Behavior relating to nutrition of sixty-six nursery school children. Doctor's thesis from the Graduate School of Cornell University, 1939, pp. 1-332.

(11) McCay, Jeanette B., Waring, Ethel B. and Bull, Helen D.: Health and development of a group of nursery school children. Child Develop., 1940, 11, 127-141.

(12) McCay, Jeanette B., Waring, Ethel B., and Kruse, Paul J.: Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters. Genet. Psychol. Monog., 1940, 22, 491-555.

(13) Murlin, J. R.: Energy metabolism. In Luck, J. M., and Hall, V. E. (eds.) Ann. Rev. Physiol., 1939, 131-163.

(14) Olson, W. C.: The measurement of nervous habits in normal children. 1929, University of Minnesota Press, Minneapolis, Minnesota, pp. xii-94.

FACTORS ASSOCIATED WITH CRYING IN YOUNG CHILDREN IN THE NURSERY SCHOOL AND THE HOME

CATHERINE LANDRETH[1]

In a recent article (8) a comparison was presented of different methods employed in the investigation of crying in nursery school children. Briefly, the results indicated that the four methods employed - time sampling, incident sampling, teachers' records, and teachers' ratings were not equally valid in furnishing a measure of the incidence of this behavior; and that, in general, selection of method should be based on the characteristics of the particular behavior under investigation. Low measures of consistency further indicated that during the eight weeks of this study different factors were operating to modify the behavior investigated in individual children.

The present article is concerned with an analysis of these factors. The purpose of this analysis is to determine

1) the extent to which physiological and environmental factors affected the behavior investigated;

2) the extent of variation in the home and nursery school situations in which this behavior arose.

As most studies on behavior of young children are made in a nursery school, it is of interest to determine differences in home and nursery school situations and differences in children's responses in these situations which might indicate the extent to which nursery school behavior is representative of home behavior. Data for this analysis were furnished from the following daily record sheets made over a period of eight weeks in the nursery school, and similar incident sampling records made over a period of five weeks in the children's homes. Home records included a daily report of variations in the child's physical condition and daily routine.

Records in the nursery school were made on thirty-two children, fourteen girls and eighteen boys with an age range at the midpoint of the investigation from two years, eight months to five years, two months. The children had been in attendance at the Institute of Child Welfare nursery school for periods ranging from 36 to 393 days. Mean and Median IQ's on the California Preschool Scale were 122 and 124 respectively. The group represented a selected socio-economic sample in that they were drawn from middle class homes of professional and business people. The selection arose out of the need for obtaining intelligent cooperation from the parents in the matter of record keeping. The home records were obtained on twenty-five of the thirty-two children, eight girls and seventeen boys.

As indicated in an earlier publication (8) the incident sampling method provided the most valid measure of the children's crying behavior in the nursery school. Reliability of the observations expressed in terms of percentage agreement between two independent observers ranged from 100 per cent to 87 per cent with a mean of 95 per cent.

[1]From Institute of Child Welfare, University of California.

CHILD DEVELOPMENT, Vol. 12, No. 2 (June, 1941)

OBSERVATIONAL STUDY OF CRYING
Daily Record Sheet

(Use a fresh sheet every day even if there are no episodes involving crying)

Date_____ Child's Name_____

Physical condition (check): ....Normal ....Heavy Cold

....Slight Cold....Digestive Upset

....Other Sickness (indicate)_____

_____

Check any of the following occurring during the day:

.....Adult visitors in the home .....Child taken out driving

.....Child visitors in the home .....Child given new toy

.....Child taken visiting or calling .....Child given new food

.....Child taken shopping

Other events outside daily regime _____
(Such as going to doctor or dentist, to church, to movies, etc.) _____

Were adults in home under any unusual strain from sickness or other causes? (check).....

Health Routine:

Child in bed previous night at _____
asleep at _____
awake at _____
up at _____

Night sleep sound_____ or restless_____
Bed dry_____ or wet_____

Day nap from_____ to_____
Bed dry_____ or wet_____

Bowel movements at_____
Movements normal? Yes_____ No_____

Breakfast at_____
Lunch at_____
Supper at_____
Food between meals? At_____

Hours not included: From_____ to_____

OBSERVATIONAL STUDY OF CRYING
Daily Record Sheet (continued)

List here each instance of crying with tears, screaming, sobbing, whining, whimpering, or 'squealing' occurring in the preceding twenty-four hours.

		Incidents 1.	2.	3.
TIME	Time of occurrence			
	Duration of outburst			
	Reported by			
SITUATION	What was child doing at the time?			
	Who were present?			
	What 'set off' crying?			
CRYING BEHAVIOR	Crying with tears			
	Screaming			
	Sobbing			
	Whimpering			
	Whining			
	'Squealing'			
ACCOMPANYING BEHAVIOR	Words spoken			
	Limpness			
	Rigidity			
	Kicking or stamping			
	Hitting or striking			
	Throwing self on floor			
	Holding the breath			
	Other			
BEHAVIOR AFTER OUTBURST	Cheerful			
	Fretful and unhappy			
	Other			
REACTION OF WHOM PRESENT	Type of response			
	By whom?			
EFFECT ON CHILD'S BEHAVIOR	Stops behavior			
	No effect (continues)			
	Increase			

FACTORS RELATED TO INCIDENCE OF CRYING

Apart from the immediate situations causing crying there are several factors which in popular opinion are considered to affect the ease and frequency with which a child becomes emotionally upset. Slight colds, digestive upsets, lack of sufficient sleep, constipation, et cetera, are among those most frequently cited. As a fairly complete record summarized in Table 1 of each child's health routine for five weeks was obtained from the parents, it was possible to determine the relationship between each factor and the crying behavior of the children in this study.

As percentage of time spent crying was found to give the most representative measure of children's crying behavior at home and at school, all correlations in Table 2 were obtained between this measure and the factors listed.

Chronological Age, Length of Nursery School Attendance

For this particular group of children, chronological age and length of nursery school attendance had no significant relationship with percentage of time spent crying, either at home or at school.

Bladder Control

Irregularities of bladder control were also unrelated to the percentage of time children spent crying at home or at school.

Intelligence Quotients

There was a slight tendency for children with high IQ's to cry less than those whose IQ's were lower, particularly in home situations.

Slight Colds and Minor Defective Health Conditions

Slight colds and minor defective health conditions were associated with an increased amount of crying at home. That no significant relationship was obtained between this factor and crying in the nursery school is probably explained by the fact that children were not in the nursery school when their mothers considered their physical condition not normal.

Hours of sleep

The tendency for children who slept more to cry more is consistent with Goodenough's (4) low positive correlations between average hours of sleep and frequency of anger outbursts. In view of the low correlations between chronological age and percentages of time spent crying, age can hardly be considered a partial factor in this positive relationship between crying and hours of sleep. Rather, it would seem possible that a child who habitually sleeps somewhat more than the average may be suffering from a mild toxic condition which in turn may affect his emotional behavior. Further, the strain imposed on a child who has several emotional upsets during the day may make the additional time spent in sleeping a necessity for adequate recuperation.

TABLE 1

VARIATIONS IN PHYSICAL CONDITION
AND HEALTH ROUTINE OF CHILDREN

Case No.	% times physical condition "normal"	Departures from routine per day average	Hours sleep per day average	% days bowel movement	% days bed wet	% time spent crying home	% time spent crying school
4	100.0	0.9	12.6	48.6	0.0	1.09	1.09
10	100.0	1.3	11.3	97.1	0.0	0.20	0.00
17	96.4	1.5	13.7	92.9	3.6	0.39	0.17
18	94.3	1.7	12.7	97.1	2.9	0.08	0.19
5	94.3	1.1	12.1	65.7	0.0	0.03	0.33
15	91.4	1.2	13.0	80.0	31.4	0.17	0.05
26	91.4	0.7	11.2	91.4	0.0	0.20	0.19
33	91.4	1.5	11.9	80.0	0.0	0.40	0.35
25	88.6	1.3	12.5	100.0	5.7	0.73	0.22
30	88.6	1.2	11.8	100.0	0.0	0.28	0.38
21	85.7	0.7	13.7	100.0	0.0	1.18	1.02
24	80.6	1.0	11.7	80.6	19.4	0.04	0.25
6	80.0	0.3	11.9	85.7	51.4	0.17	0.00
14	80.0	1.1	11.6	91.4	0.0	0.14	0.02
19	80.0	0.6	11.6	88.6	3.3	0.79	0.45
28	74.3	1.2	11.4	97.1	0.0	0.30	0.14
9	74.3	1.4	12.7	88.6	0.0	0.07	0.16
8	74.3	1.0	12.2	94.3	0.0	0.15	0.13
16	74.3	1.5	13.2	97.1	91.4	0.31	0.20
22	71.4	0.6	11.8	100.0	31.4	0.60	0.19
23	71.4	0.8	11.6	80.0	0.0	0.21	0.09
31	71.4	0.6	12.0	94.3	25.7	0.34	0.36
13	62.9	0.8	12.5	91.4	0.0	1.11	0.42
7	54.3	0.5	12.9	85.7	71.4	0.94	0.26
29	51.4	0.5	13.8	85.7	11.8	2.02	0.98
Mean	80.9	1.0	12.3	88.5	13.9	0.48	0.31
Median	80.9	1.0	12.1	91.4	0.0	0.30	0.20
Range	51.4	0.3	11.2	48.6	0.0	0.03	0.00
	100.0	1.7	13.8	100.0	91.4	2.02	1.09

TABLE 2

FACTORS ASSOCIATED WITH CRYING BEHAVIOR

Factors	% Time spent crying at home	% Time spent crying at school
Chronological Age	-.18	-.11
Intelligence Quotient (California Preschool Mental Test Scale Form A)	-.38	-.29
Length of Nursery School Attendance	----------	.05
Physical condition (% of days condition normal, by mother's report)	-.51	-.12
Average hours sleep	.53	.49
Bowel Movements (% of days regular movement)	-.13	-.39
Bladder control (% of days bed wet)	.01	-.20
Departures from daily regime (Averages for adult and child visitors in home, simple diversions, etc.)	-.49	-.31

Regular Elimination

Evidence in regard to a relationship between incidence of crying and regularity of bowel movements is inconclusive; the correlations between this factor and crying behavior in the home and school are .13 and .39 respectively.

Departures from Daily Routine

While at first sight it would seem that a certain amount of diversion and variety in the daily routine, so long as it does not interfere with the child's health program, actually favors fewer emotional upsets, this may not be the case. It is equally possible that mothers whose children are less prone to emotional upsets are more apt to take them out with them. Child 18, who has the fewest incidents and spends the least percentage of time crying at home has the highest average number (1.7) of departures from routine per day, while child 29, who has the highest score for incidence of crying at home, has the second lowest average number of departures from routine. That this finding is not in agreement with Goodenough's (4) may be due to a difference in the type of departures recorded and also a difference in the age range of the group.

Sex Difference

The means for percentage of time spent crying at home for boys was .57 per cent as compared with .26 per cent for the girls, the critical ratio based on standard deviations being 2.04. Means for percentage of time spent crying at school for boys and girls were .34 per cent and .27 per cent respectively, with a critical ratio of .77.

The tendency for boys to cry more than girls was more apparent in the home than the school. The difference on the proportion of boys in the home and school samples was doubtless a modifying factor. It is possible that boys of this socio-economic group are more irked than girls by complying with routines and parental restrictions of activity. Goodenough's (5) findings on sex and socio-economic differences on negativism in nursery school children during test situations would seem to support such a view.

Day of Week

The number of crying incidents per child per day present in the nursery school varied from 1.2 to 1.4 and would seem to indicate that the day of the week was not a significant factor in incidence of crying.

Hour of Day

The percentage of total incidents of crying occurring between each of the six half hour periods of the school morning, 9 A.M. to 12 A.M., are presented in Figure 1. During the first half hour 9:00 to 9:30 the children were still in process of arriving so that actually fewer children were present for this period. The last half hour, 11:30 to 12:00, was devoted to the putting away of toys and the story period. In the four periods from 9:30 to 11:30, which are comparable in terms of number of children present and type of activity engaged in, the last has the greatest percentage of incidence, indicating a rise toward the end

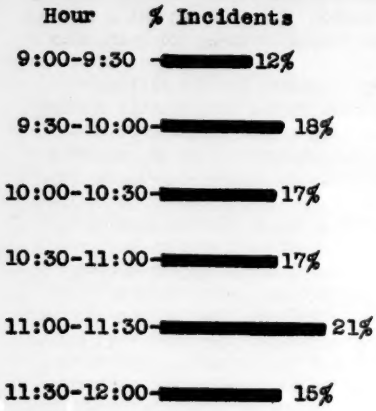


Fig. 1. Relation of incidence of crying in the nursery school to hour of day.

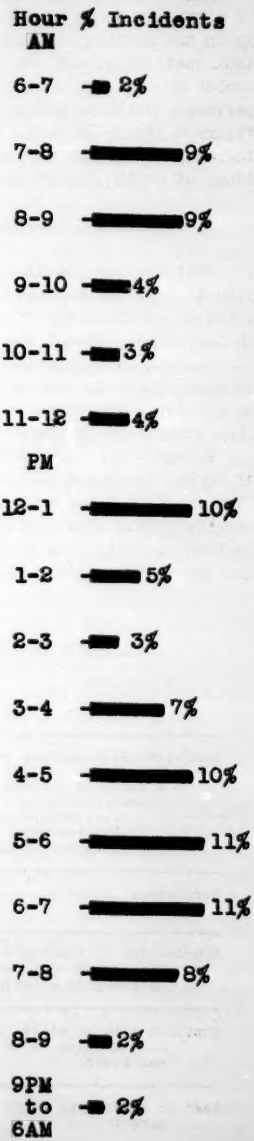


Fig. 2. Relation of incidence of crying in the home to hour of day.

of the morning.

Many factors undoubtedly led to differences in the amount of crying at different hours of the day in the homes. The time the children get up in the morning, the hours they were away at school, the hours they took their naps, and the time they went to bed necessarily affected the number of children on whom records were obtained for the different periods. While no analysis was undertaken of the many details involved, Figure 2 simply presents the situation with regard to the percent of incidents occurring during the day, and indicates a rise in the incidence of crying toward the end of the day.

IMMEDIATE SITUATIONS CAUSING CRYING IN THE NURSERY SCHOOL

What are the immediate situations causing crying in the nursery school? Are sex and age determining factors in regard to these situations? Is there any similarity between the situations causing crying in the nursery school and those causing crying at home?

In this investigation it was found that the immediate situations causing crying in the nursery school could be classified into seven categories. The percentage distribution of crying incidents resulting from situations in these categories is presented in Table 3.

Though there were marked individual differences in the distribution of crying incidents in these seven categories, the mean and median figures of 70.3 per cent and 76 per cent respectively indicated that conflicts with other children were responsible for the majority of incidents. Only five children had less than fifty per cent of their incidents attributable to this cause and only one child, the second

TABLE 3

IMMEDIATE SITUATIONS CAUSING CRYING IN THE NURSERY SCHOOL

Situations	Percentage distribution of 1245 crying incidents
Conflicts with another child: Attacks on child's person; Attacks on child's property; Frustration by another child i.e. other child would not let him play with him or use his play materials	75.7
Accidental injury	9.1
Frustration by inanimate objects, i.e. child could not make one of the play materials work	8.1
Conflict with an adult, i.e. child did not wish to do what he was asked	4.9
Fear or insecurity as in leaving parent	1.6

youngest in the school, had no instances of crying because of conflicts with other children. As he played alone during practically the entire period of observation the lack of conflict was doubtless due to his lack of contact with other children.

Accidental injury was a much less significant cause of crying, with mean and median figures for the group 13.7 per cent and 13.1 per cent respectively. Only three children never cried from this cause. Mean and median figures for frustration by inanimate objects were 8.5 per cent and 1.7 per cent respectively, with fifteen of the thirty-two children having no crying incidents associated with this situation.

Only twelve of the thirty-two children were recorded as crying because of conflicts with adults in the nursery school. A total of sixty-two incidents, less than five per cent resulted from this cause; whereas in the homes, twenty-one per cent of the total incidents reported were directly caused by conflicts with adults and a large percentage of the remainder doubtlessly indirectly attributable to the same cause. It is of interest that child 29, whose record shows the highest percentage of crying due to conflicts with adults in the school, has a home record which is also the highest. As this child's intelligence quotient was only 86 and his language development somewhat retarded, it is possible that these were significant factors for his behavior in both situations.

Twenty crying incidents were attributed to fear or "insecurity" on the part of the seven children for whom they were recorded. In the main these were associated with leaving a parent on arrival at the nursery school.

Sex Differences in Relation to Types of Nursery School Situations Which Caused Crying

Figure 3 presents mean percentage distribution of crying incidents for boys and girls in the seven categories listed. In the case of crying incidents arising from conflicts with other children, the obtained means, (in terms of percentage) tend to be higher for girls. However, since girls have fewer total crying incidents, the actual number of incidents due to conflicts is less in the case of girls (394) than boys (549).

A sex difference is apparently present in the category of accidental injury. Not only is the mean for girls higher, but the actual number of incidents for the girls is 75 against the boys' 44. Either the girls lack the skill and coordination to prevent as many tumbles and mishaps as the boys do, or crying because of injury receives more social disapproval and is more often inhibited in the case of boys than girls.

As Figure 3 indicates, the boys had a higher average per cent of crying incidents arising from frustration by inanimate objects than the girls. Further, eighty-eight of the one hundred and one incidents of crying resulting from this cause were recorded for boys. This is possibly due to the greater activity of the boys and the type of play materials they used, a sex difference that was observed incidentally during this study and noted by Van Alstyne (9). It would also seem from the means for situations caused by conflicts with adults that the girls as a group were somewhat less resistant to adult authority. In

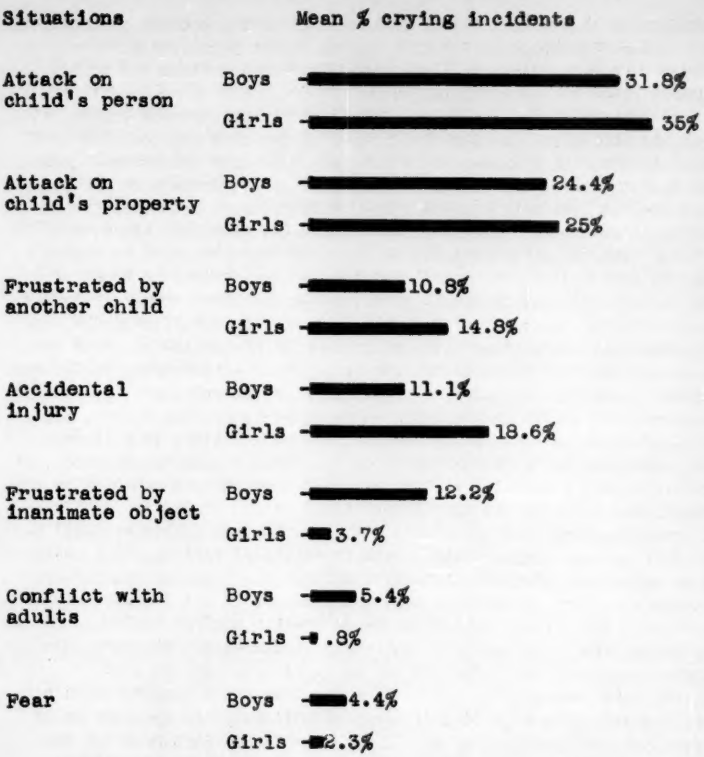


Fig. 3. Immediate situations causing crying in the nursery school.

terms of incidents, fifty-eight of the sixty-two resulting from conflict with adults were recorded for boys, a finding consistent with Goodenough's (5) and Caille's (2) studies on negativism in nursery school children.

The difference between the percentage of crying incidents for boys and girls resulting from fear is doubtless attributable to the record of one boy who had 45 per cent of his crying arising from this cause. Actually, eleven of the twenty incidents of crying resulting from fear were recorded for girls.

In summary, apparent sex differences for this group were the girls' greater percentage frequency of crying because of injury and the boys' greater percentage frequency of crying because of frustration by inanimate objects and conflicts with adults.

Analysis of Children's Conflicts

As seventy-five per cent of the total incidents of crying in the

nursery school resulted from conflicts with other children, the extent to which particular children were associated with, if not directly responsible for, crying in other children seemed a fruitful study for further investigation. The first fact which became apparent was the marked difference in the frequency with which boys and girls were associated with other children's crying. In 943 incidents of crying because of conflict with another child, the "other child" was a boy in 765 of the incidents and a girl in only 178. By reducing the total number of incidents in which a child was associated with another's crying to the average per child present, results for different children and for boys and girls were made directly comparable. The average per day per girl was .41 of an incident against 1.31 incident per boy per day, a ratio of roughly one to three. This difference would seem to be due to greater aggressiveness on the part of the boys or to a sex difference in language development leading to greater use by boys or direct action in preference to speech. A sex difference in regard to conflict behavior has previously been reported by Jersild (7), Dawe (3), and Green (6).

Apart from the sex differences there were considerable individual differences not only in the number of incidents of crying children were associated with, but also in the particular children with whom they had conflicts. Child 29 was the "other child" in 182 crying incidents; 121 of these incidents were conflicts with his constant playmate, child 21. Child 21 was the "other child" in 82 incidents of crying, but in only 21 of these was number 29 the child crying.

The situation concerning these two children is interesting. Child 21 entered the nursery school unable to speak a word of English; he is the son of French parents who speak only French to him at home. His language handicap was reflected in his IQ of 101. At the time he entered child 29 had very little language also. His language handicap, however, seemed to be definitely associated with retarded mental development as evidenced by his IQ of 86, obtained at his second and most recent mental test. In his social responses to other children number 29 was definitely immature. The friendship between these two children, who habitually played together, would seem to have sprung originally from their common lack of language. The fact that this lack of language resulted from different causes is one reason why the close association between the two children was a tearful one for number 21. On the days on which number 29 was absent, number 21 either did not cry at all or cried much less frequently. While it is true that this child ranked first for duration of crying at home and second at school, the crying at home was based on fewer incidents. Further, there are indications that the crying in the nursery school may have actually affected behavior at home. After a morning in which number 21 had sixteen incidents of crying in school, his mother reported on his home record at noon, "cross and fretful for no apparent reason."

Another record which is of interest is that of child 19. Of the forty-two incidents of crying he was associated with, twenty-seven were with his sister, (child 20). Of her fifty-five incidents of crying arising from conflict with other children, one half were associated with her brother. She in turn was associated with only six of

her brother's twenty-nine incidents of crying, in conflict with other children. At home as in school, number 19 apparently spent a great part of the time in conflict with his sister despite the fact that his mother asserted that he had never been jealous of her.

Child 10, though associated with 127 incidents of crying, cried only once himself in conflict with another child. This child, a boy, was the oldest in the nursery school, and much larger and stronger than the rest of the children. Despite his pleasant friendly disposition, the fact that he had really outgrown the nursery school activities was largely responsible for the fact that he tended to dominate the activities of the younger children.

IMMEDIATE SITUATIONS CAUSING CRYING IN THE HOME

The immediate situations causing crying in the home were found to lend themselves to classification under seventeen categories. Seven of these were concerned with the establishment of such routine habits as:

Sleeping
Taking naps
Elimination
Washing and bathing
Brushing and washing hair
Dressing and undressing
Eating

Five were associated with play activities. They were listed as:

Frustration from play materials
Frustration from pets
Conflict with siblings
Conflict with other children
Conflict with parents or adults in the home

Five which came under no general heading were:

Injury and treatment
Medical and dental examinations
Insecurity
Fussiness
Coming to the Nursery School

As the classification into separate mutually exclusive categories permitted the possibility of a certain amount of subjective interpretation on the part of the investigator, a nursery school teacher also classified one hundred and five of the one thousand and thirty seven home records in regard to both the situation causing crying and the type of response which the child received. Disagreement between the two classifications occurred in only six per cent of the situations, indicating that the classification was reasonably reliable.

As the seventeen categories listed would seem to cover practically all the situations occurring during the day, somewhat more significance attaches to the frequency with which particular situations provoked emotional upsets. A summary of both the situations and frequencies is presented in Figure 4.

The parents' record of the situations furnished evidence that some

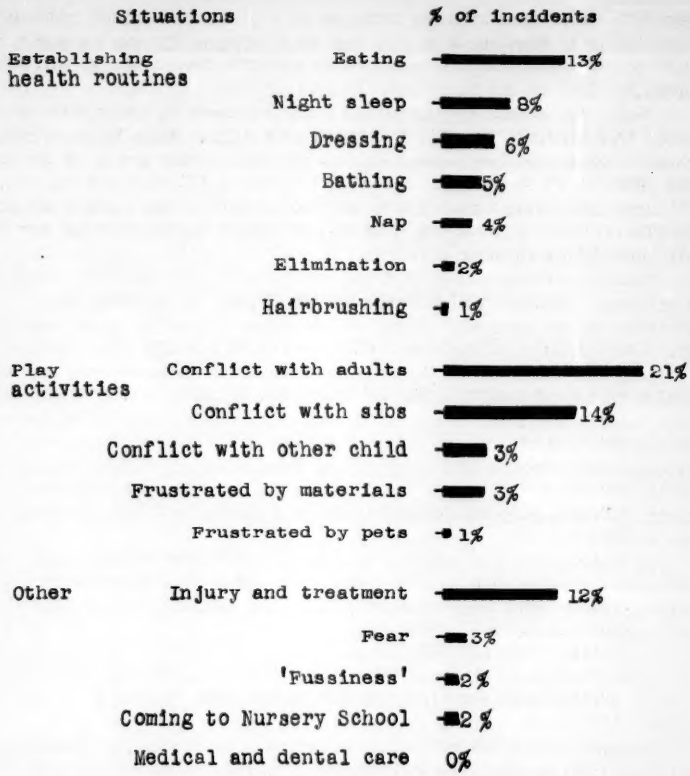


Fig. 4. Immediate situations causing crying in the home.

of the difficulties arose out of inconsistencies and poor techniques in child training.

For example, some of the children who objected to going to bed had variations in their bed-hour of from one to two hours, an irregularity that would hardly make a regular bed-hour acceptable. Child 6, forty-two per cent of whose crying was concerned with either going to bed or getting up, had a sleeping schedule which seemed unsuited to him. This child was usually given a three or occasionally four hour nap ending at four-thirty and sometimes five. Seven-thirty found him unwilling to go to bed again, clamoring for toys to take to bed with him and taking an hour or more in actually getting to sleep. On being awakened at six or six-thirty after nine or nine and a half hours sleep, he was apparently unrefreshed and cried over the necessity of getting up.

The most frequent difficulty in connection with dressing and undressing occurred over the children's wishing to wear particular

garments while their mothers insisted on a different choice. Typical is the case of a four-year-old girl who cried fifteen minutes to wear a pair of red socks and then appeared at school red-eyed but in yellow socks.

Under the heading Eating Habits sixty per cent of the difficulties arose through conflict with the adults over eating meals or particular foods: forty per cent represented conflicts over the manner of eating and behavior at the table.

Only one child, number 17, cried over coming to the nursery school. Similarly, his crying in the nursery school was almost entirely due to his not wishing to stay.

Conflict with siblings occurred in every case in which the child had a sibling. The particular situations would seem to indicate that jealousy was an important factor in the cases of number 19 and number 25.

Only one child had no crying incidents resulting from conflict with adults in the household. Many of the difficulties encountered under this heading would also seem to reflect injudicious handling on the part of parents. Situations such as "Parent spanks child for lying over misdemeanor", "Parent spanks child for striking him", "Parent spanks child for calling him names", are not suggestive of a particularly desirable parent-child relationship on which to base mutual respect and esteem.

In fairness to the parents, it must be added that though their records frequently indicated the use of unsatisfactory methods, the very nature of this study tended to make it a record of their less successful methods of child training.

RESPONSES FROM ADULTS AND CHILDREN TO SITUATIONS IN WHICH CRYING OCCURRED IN THE NURSERY SCHOOL

Bridges (1) has stated in connection with the crying of preschool children that it pays to cry. The findings from this investigation would lead one to believe that in this nursery school there was a law of diminishing returns operating for those who cried more often. Figure 5, which summarizes the percentage of responses to crying incidents in the nursery school, shows that number 21, who had the greatest number of crying incidents in the nursery school, received an adult response 63 per cent of the time. Sixty-one per cent of the children received a greater percentage of responses than he did. Child 4, who had the next largest number of incidents, received a response fifty-two per cent of the time. Ninety-four per cent of the children received a greater percentage than this. Of the ten children who ranked highest in respect to number of incidents of crying, only one received more than the median percentage of responses.

In regard to the type of response, adult responses were classified in eight categories represented in Figure 6 in terms of percentage frequency. The responses from children were similarly classified, save that the category "Ridiculing the child crying" was substituted for "Isolating child crying". It will be seen in Figure 5 that the children numbers 21, 29, and 4, who cried most often, were among the few who were

Fig. 5. Adult responses to children crying in the nursery school.

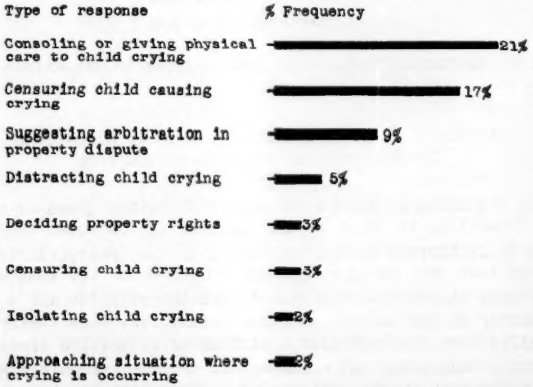


Fig. 6. Adult responses to situations causing crying in the nursery school.

censured or isolated for crying.

The most frequent response of the children to crying was consoling the child crying. Next in order came approaching the situation, ridiculing the child crying, and censuring the child causing crying. The three children numbers 21, 29, and 4 received the greatest percentage of ridicule from the children. In the light of these findings, it is difficult to believe that these children derived much satisfaction from the social environment through their crying.

Fig. 7. Responses from adults and children to situations causing crying in the home.

RESPONSES FROM ADULTS AND CHILDREN TO SITUATIONS IN WHICH CRYING OCCURRED IN THE HOME

Figure 7 summarized the percentage of different types of responses given to situations in which crying occurred in the home. While the selection of categories was largely based on the interpretation on one person, the fact that an agreement of ninety-four per cent was obtained between classifications made by the investigator and a nursery school teacher on one hundred and five crying incidents indicated that the classification was made with a minimum of subjective error.

Ignoring, reasoning, and spanking, in order of frequency were apparently the most favored methods of dealing with the type of situations which led to crying, though considerable individual differences were apparent in parents' methods. Twenty-four per cent of parents never resorted to spanking, while two recorded spanking in twenty per cent of the incidents reported. Only two of the twenty-five parents did not report using ignoring or reasoning as methods of coping with the situations encountered.

As the type of situation in which crying occurred necessarily influenced to some extent the type of response, a summary was made of the types of responses given to different situations. In decreasing order of frequency spanking was most often used in situations involving: (a) conflict with the parents, (b) sleeping, and (c) taking naps; these apparently represent some sort of parental breaking point. In seven of the seventeen situations listed it was never used.

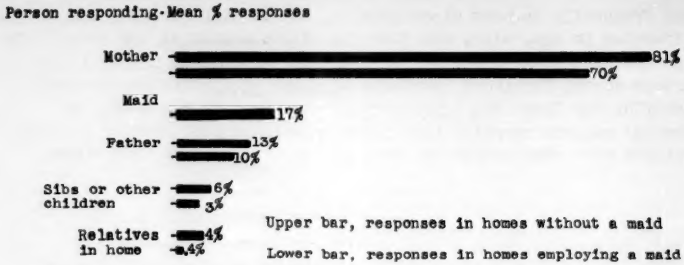


Fig. 8. Responses from adults and children to situations causing crying in the home.

Reasoning was used at some time by parents in all situations, and ignoring was used in all save two of the situations. Apart from the rather general use of these two methods, there were indications, as illustrated by spanking frequencies in different situations, that the methods parents used in handling problems were partly dependent on the nature of the problem and not representative of any particular bias toward the efficacy of any one method. As Figure 8 indicates, the mother in the home responded to the great majority of situations in which crying occurred. The fact that the maid, in some cases only in the home for a few hours a day, responded to a greater percentage of situations than the father is not without its implications for the type of training required for maid service.

SUMMARY

This study was concerned with an analysis of the factors associated with crying in young children in the nursery school and the home. It sought specifically to determine

1) the extent to which physiological and environmental factors affected the behavior investigated;

2) the extent of variation in the home and nursery school situations in which this behavior arose.

Eight weeks of incident sampling records on thirty-two children in the nursery school, and five weeks incident sampling and daily routine records on twenty-five of these children in their homes furnished the data for the investigation.

The results revealed significant relationships between some of the factors investigated and the incidence of crying behavior. These relationships furnished an explanation for the low measures of consistency obtained for this behavior. The type of situations in which crying occurred in the nursery school differed from those in the home, the majority of incidents in the school being associated with conflicts with other children, and in the home with the carrying out of health routines and conflicts with the parents. An analysis of the children's conflicts in the nursery school revealed that boys were much more frequently associated with other children's crying than girls, the frequency ratio being three to one. There were also indications of conflicts occurring

more frequently between close friends, and arising out of a marked difference in age, size, and strength of one member of the group. The type and percentage frequency of response to children crying in the nursery school furnished evidence of a law of diminishing returns operating for those who cried more frequently. In the homes the parents' records revealed that inconsistent and poor methods of child training were responsible for many of the difficulties that arose.

REFERENCES

(1) Bridges, K. M. B.: The social and emotional development of the preschool child. London, K. Paul, Trench, Trubner, 1931.

(2) Caille, R. K.: Resistant behavior of preschool children. Child Development Monographs, 1933. No. 11.

(3) Dawe, H. C.: An analysis of two hundred quarrels of preschool children. Child Develop., 1934, 5, 139-157.

(4) Goodenough, Florence L.: Anger in young children. Minneapolis, Univ. Minnesota Press, 1931.

(5) Goodenough, Florence: The emotional behavior of young children during mental tests. J. Juv. Research, 1929, 13, 204-219.

(6) Green, E. H.: Group play and quarreling among preschool children. Child Develop., 1933, 4, 302-307.

(7) Jersild, A. T. and Markey F. V.: Conflicts between preschool children. Child Development Monographs, 1935, No. 21.

(8) Landreth, Catherine: Consistency of four methods of measuring one type of sporadic emotional behavior (crying) in nursery school children. J. Genet. Psychol., 1940, 57, 101-118.

(9) Van Alstyne, Dorothy: Play behavior and choice of play materials of preschool children. Chicago, Univ.Chicago Press, 1932..

VOWEL ELEMENTS IN THE CRYING VOCALIZATION OF INFANTS UNDER TEN DAYS OF AGE

ORVIS C. IRWIN[1] AND THAYER CURRY[2]

The present status of research in initial speech sound equipment of the infant has been presented elsewhere (2). It is indicated that there does not exist a large body of data secured from adequate samplings of infants for purposes of statistical analysis, that most of the observations in any given study were made on one or two children, that usually no systematical research methods were formulated and employed, that statistical techniques essential to the analysis of mass data are practically absent, that reliabilities of observers seldom have been established, and that most observers have used an alphabetical rather than a phonetic system of symbols for recording.

The purpose of the present endeavor is to systematically study the vowel equipment during crying of infants under ten days of age. Forty infants varying from one to ten days in age were placed in a Pratt experimental cabinet and their crying sounds were independently transcribed in the International Phonetic Alphabet by two observers. A broad transcription of this alphabet as suggested by Fairbanks (1) was used.

RELIABILITY

The problem of reliability of observers is fundamental to this work. A first attempt involved the use of the short time sample. By means of a stop watch all sounds occurring during one minute were recorded. It was found that this method yielded agreement between the records little better than afforded by chance. Therefore it was abandoned in favor of a method which employed as a standard observation an easily identified behavior unit whose duration is well within the ordinary attention span. A behavior unit which meets this requirement is the single respiration or breath. Since babies vocalize upon the phase of exhalation the problem of the recorder is to focus his attention upon a given breath and record such sounds as occur upon it. In establishing reliabilities of two transcribers, one of them indicates, by pointing, a respiration about to take place as the next one to be observed and recorded. The attention of both recorders is then focussed upon it and such sounds as are carried by it. Consecutive respirations or every other respiration, or any other type of sample, can be observed quite readily. In this study most of the sounds were secured from consecutive exhalations, although a few were from spaced respirations.

A record of the sounds occurring on twenty-five respirations of each of forty infants was made by the two observers. This yielded a sample of 1,000 respirations. A large majority of the respirations carried single sounds, about 200 of them carried more than one sound, and only a very few as many as four sounds.

[1]*From Iowa Child Welfare Research Station, University of Iowa.*
[2]*From University of Oklahoma.*

CHILD DEVELOPMENT, Vol. 12, No. 2 (June, 1941)

The reliability between observers is indicated by the following analyses. The per cent of agreement was determined for each infant and the average per cent for the forty babies was found to be 85 per cent. The following tabulation indicates the degree of success of the method with this type of data.

Agreement, Per Cent	Number of Infants
100	7
96	8
92	5
88	6
84	2
80	3
72	1
60	1
56	5
52	2
Total	40

The tabulation reveals that with seven infants the transcriptions were identical, that with twenty infants the per cent of agreement is above 90, with thirty-one babies it is above 80, and with nine infants the agreement is not satisfactory, lying below 80. The average of 85 per cent agreement, however, is a value comparable with much work involving the observational technique.

It would be expected that the problem of transcription of sounds occurring singly on a respiration would be less difficult than when two or more sounds occur. About 800 respirations, a considerable majority, carried single sounds. An analysis of these sounds reveals the degree of agreement between transcriptions. The following tabulation gives the number of each of the five vowel sounds heard singly by the two observers, the per cent of the total for each sound, and the per cent of agreement for the most frequently used vowels.

Vowel	Frequency		Frequency, Per Cent		Agreement, Per Cent
	C	I	C	I	
æ	710	746	90.2	89.7	95
ɛ	63	68	8.0	8.2	93
ʌ	7	6	.9	.7	
ɪ	5	7	.6	.8	
i	2	5	.3	.6	
Total	787	832	100.0	100.0	

The tabulation shows that observer C recorded the vowel æ 710 in a total of 787 vowels recorded as single sounds by him, yielding a per cent of 90.2. Observer I heard this sound 746 times in a total of 832 sounds. This is 89.7 per cent, quite comparable with the value of 90.2

per cent. The agreement between these observers in recording the single vowel æ is 95 per cent. In the case of the vowel ɛ, it is 93 per cent.

It should be emphasized that the above work was done on newborn infants in whom presumably the speech sound equipment is more meager than with older children. There is also the problem of reliability of the method with children whose speech sound equipment is more extensive.

In an unpublished study (3) on the vowel equipment of a group of infants during the second quarter of the first year of life, the reliability of two observers using the behavior unit method was calculated on data from five subjects. The tabulation below shows the number of times observer I heard eleven vowels in a total of 487 non-crying utterances, and observer K in a total of 490, while both were transcribing simultaneously and independently. The record included thirty respirations of each of the five subjects. Approximately 490 vowel sounds and 300 consonant sounds occurred in a total of 150 breaths. The third column gives the percentage of agreement between observers.

Vowel	Observers I	Observers K	Agreement, Per Cent
i	42	37	88
ɪ	57	60	95
e	7	5	71
ɛ	69	77	90
æ	55	57	96
ʌ	85	92	92
a	7	8	88
ɒ	23	26	88
o	12	12	100
ʊ	122	108	89
u	8	8	100
Total	487	490	

The tabulation shows that all but one of the values are above 80 per cent, the range is from 71 to 100 per cent, and the mean per cent of agreement is 91. When the Chi square test of homogeneity is applied to the two columns of frequencies the probability value is very much smaller than 1 per cent, indicating substantial agreement.

The vocalization of consonant elements is negligible in crying newborn infants. However, the method has been tested on non-crying consonant sounds of the five older infants. The tabulation shows good agreement between observers.

It is seen that all the percentages except one lie above 80, and eleven of the thirteen above 90. The average is 94 per cent. Here also the test for homogeneity of the two arrays gives a probability value of less than 1 per cent, indicating substantial agreement.

In general, then, it may be said that the use of the unit behavior method for the transcription of the phonetic elements in the speech of young infants yields a satisfactory degree of agreement between trained observers.

Consonants	Observers I	Observers K	Agreement, Per Cent
h	150	148	99
g	11	11	100
k	38	37	99
t	27	27	100
p	6	6	100
b	18	17	95
d	28	27	96
m	3	3	100
n	5	5	100
x	2	2	100
j	10	8	80
v	2	1	50
l	3	3	100
Total	303	295	

RESULTS

The remaining part of this report is concerned with findings obtained by this method with the vowel vocalization of forty infants during the first ten days of life. The transcriptions of observer C have been used for this purpose.

Table 1 gives the number of infants using the vowel elements and the total number of times these sounds were used. This table reveals that only four vowel sounds are used by newborns with any degree of frequency. These in order are æ, ɛ, ʌ, and I. It is seen that the only vowel element used by all forty babies is the æ sound. The ɛ sound is uttered by 34 infants, and I and ʌ by 21 and 17 subjects respectively. The remaining vowels are used by few of the children. Moreover it is evident that the element æ strikingly exceeds all other vowel sounds in frequency. Data for non-crying sounds have not been included because of the fewness of these sounds in our transcriptions. Newborn infants rarely utter non-crying sounds and the collection of an adequate sample requires a tremendous amount of time and patience.

Table 1

Number of Infants Using a Vowel Sound and the Frequency
That the Sound Occurred in Twenty-Five Observations
(40 Neonates)

Vowels	i	I	e	ɛ	æ	ʌ	ə	ɑ	ɒ	ɔ	o	ʊ	u
Number of cases	6	21	1	34	40	17	0	0	0	1	0	3	1
Frequency	14	72	1	195	903	89	0	0	0	1	0	8	1
Mean frequency per case	2	3	1	6	23	5	0	0	0	1	0	3	1

The dispersion of the use of the vowel elements for each of the forty infants is presented in Table 2. It is arranged into the three phonetic categories of the vowel elements, front, middle, and back vowels, indicating the mouth parts used in the production of these sounds.

Aside from the piling up of the ɘl sounds, the table is noteworthy for the fact that of the three categories the front vowels clearly are the most frequent. One of the middle vowels is appreciably present but the back vowels are conspicuously infrequent.

This result has been summarized in the following tabulation:

Vowels	Frequency, Per Cent
Front	92
Middle	7
Back	1
Total	100

Front and middle vowels constitute 99 per cent of all vowels produced by newborns and back vowels only 1 per cent.

These data indicate decisively, then, that newborn infants use almost exclusively such vowels as are made by the front mouth parts and do not use the vowels formed by the back parts of the mouth. Since older children and adults employ all mouth parts in speaking, the course of development during the first and second years is indicated. It must consist in the increasing use of back vowels.

In order to render the values in Table 2 comparable, they are given as percentages in Table 3. In addition, the data have been separated according to sex. The mean percentages for each vowel for the boys and the girls lie close together. In the case of the sound ɘ, the percentage for the girls is 73, for the boys it is 67. Table 4 shows Students' t-Test to determine the significances of these differences. The t-ratios all are too small to be of consequence.

It is of interest to learn if there are differences in the use of vowel sounds from day to day during the first ten days. The situation is revealed in Tables 5 and 6. Since the number of infants for each day is rather small, the data have been recast into three periods; 1 to 4 days, 5 to 7 days, and 8 to 10 days. It is apparent from Table 7 that the differences are not statistically significant. The minimum t value at the 5 per cent level would have to be 2.06 to indicate significance. None of these ratios equal this value.

In order to compare the frequency of vowel usage in newborn and adult speech, data from Voelker's (4) study on phonetic dispersion in vowel pronunciation in radio announcements has been selected because it represents the most adequate and carefully collected materials in the literature of this subject. The tabulation presents the infant and the Voelker data together with the percentages of frequency for each of twelve vowel elements for both groups.

Table 2

FREQUENCY OF OCCURRENCE OF VOWEL SOUNDS IN NEWBORN INFANTS

Infant	Front					Middle		Back						Total
Number	i	I	e	ɛ	æ	ʌ	ə	ɑ	ɒ	ɔ	o	ʊ	u	
1	3	4		7	18	1								33
2		3		16	21	5								45
3		1		4	21	4								30
4				8	21	7								36
5	6	6		14	13									39
6	1	3		11	22	1								38
7		3		13	18									34
8		3		3	24									30
9		1		4	21	3								29
10	2	8		10	17	1								38
11		12		1	25	19								57
12		6			26	11								43
13		6		1	25	11								43
14		2		13	16	6								42
15		2		5	23	9					3			42
16					25									25
17		3		10	23	3								39
18		1		8	23									32
19		2		3	24					1				30
20				7	23									30
21					25	1					4		1	31
22				3	23									26
23				4	21	1								26
24	1			4	23									28
25		1		1	25									27
26		3		3	24									30
27				2	24									26
28					25									25
29				1	25	5					1			32
30		1		6	22	1								30
31				8	22									30
32			1		25									26
33	1	1		1	23									26
34					25									25
35				9	20									29
36				2	25									27
37				2	24									26
38				1	25									26
39				1	25									26
40				4	23									27
Total	14	72	1	195	903	89				1.	8.		1.	1,284
Means	.35	1.8	.03	4.9	22.6	2.2				.03	.2		.03	
$S.E._M$	.17	.41	-	.74	.41	.64				-	.12		-	

Table 3

Per Cent of Total that Each Vowel Sound Occurs per Infant

Infant Number	Front					Middle		Back					
	i	I	e	ɛ	æ	ʌ	ə	ɑ	ɒ	ɔ	o	ʊ	u
Newborn Girls													
2		7		35	47	11							
3		4		13	70	13							
6	2	8		29	59	2							
9		4		14	72	10							
10	5	21		26	45	3							
12		14			60	26							
13		14		2	58	26							
17		8		26	58	8							
20				23	77								
22				12	88								
24	4			14	82								
25		4		4	92								
26		10		10	80								
28					100								
30		3		21	73	3							
32			4		96								
33	4	4		4	88								
36				7	93								
38				4	96								
39				4	96								
Mean per cent	.8	5.5	.2	15.4	73	5.1							
$S.E._M$	.4	1.3	.2	2.4	3.1	1.9							
Newborn Boys													
1	9	12		21	55	3							
4				22	58	20							
5	15	15		36	34								
7	9			38	53								
8		10		10	80								
11		21		2	44	33							
14		5		43	38	14							
15		5		12	55	21					7		
16					100								
18		3		25	72								
19		7		10	80				3				
21					81	3					13	3	
23				15	81	4							
27				8	92								
29				3	78	16					3		
31				27	73								
34					100								
35				31	69								
37				8	92								
40				15	85								
Mean per cent	1.7	6.9		17	67	5.7			.2		1.2	.3	
$S.E._M$	.9	1.4		3.1	4.4	2.7			.1		.7	.1	

Table 4

Student's t-Test to Determine the Significance of the Difference in the Occurrence of Vowel Sounds in Newborn Girls and Boys (Data in the Form of Per Cent of Total Sounds of the Infant)*

Vowel Sounds	i	I	e	ɛ	æ	ʌ	ɔ	ʊ	u
Boys' mean N = 20	1.7	3.9	0.0	17.0	67.0	5.7	0.2	1.2	0.3
Girls' mean N = 20	0.8	5.5	0.2	15.4	73.0	5.1	0.0	0.0	0.0
Difference in means	0.9	1.6	0.2	3.6	5.0	0.6	0.2	1.2	0.2
S.E. diff. in means	1.0	1.9	0.2	3.9	5.4	2.8	-	0.7	-
t ratio	0.9	0.8	1.0	0.9	0.9	0.2	-	1.7	-

*For 38 degrees of freedom, which is the case here, the null hypothesis could be rejected at the 5 per cent level if the t ratio were 1.95. None of the ratios here meet that criterion.

Table 5

Per Cent of Total Vowel Sounds that a Sound Occurs for Each of the First Ten Days of Infancy

Day	N	Front					Middle		Back					
		i	I	e	ɛ	æ	ʌ	ə	ɑ	ɒ	ɔ	o	ʊ	u
1	3	0	1	1	7	84	2						4	1
2	3	5	5	0	16	74								
3	3	0	6	0	22	63	9							
4	5	1	8	0	17	66	7				1			
5	4	1	6	0	16	63	14							
6	4	1	4	0	7	88								
7	3	0	0	0	14	78	7						1	
8	5	2	8	0	24	65	1							
9	5	0	1	0	13	84	2							
10	5	0	6	0	11	72	11							

The tabulation discloses an exceedingly interesting contrast. Whereas in adult speech eleven of the twelve vowels are very appreciably present, in newborns only four vowels are produced. There are further

Vowels	Newborns	Adult
i	.1	7.4
I	5.6	20.5
e	.1	5.2
ɛ	15.2	8.0
æ	70.3	8.1
ʌ	6.9	.1
ə		17.8
a		10.9
ɔ	.1	5.6
o		5.5
ʊ	.6	4.6
u	.1	5.2
Total	99.0*	98.9

*Due to rounding, these values do not total exactly 100 per cent.

differences. The sound I is the most prominent in adult usage constituting about 20 per cent of vowel sounds. In contrast to this in newborns the element æ completely outnumbers other vowel sounds, its value amounting to 70 per cent. Moreover, there is a peculiar inversion in the use of the two middle vowels ʌ and ə. Newborns use the first but not the second. On the other hand, in adult speech ʌ is the infrequently used middle vowel while ə ranks second in use among all vowels. Finally, among newborns back vowels are practically absent but among adults these sounds constitute 32 per cent, or one-third of adult vowel usage.

Table 6

Per Cent of Total Vowel Sounds that a Sound Occurs for Each Third of the First Ten Days

Days	N	Front					Middle		Back					
		i	I	e	ɛ	æ	ʌ	ə	a	ɒ	ɔ	o	ʊ	u
1 to 4	14	2	6	-*	15	71	5				-		1	-
5 to 7	11	1	4		12	76	7						-	
8 to 10	15	1	4		16	74	5							

*Percentage too negligible to indicate.

Table 7

Student's t-Test to Determine the Significance of the Difference in the Occurrence of Vowel Sounds in Newborns for Different Thirds of the First Ten Day Period (Data in Per Cent Form)

Vowel Sounds	Days Compared	Difference in Means	S.E.Diff.	t Ratio
i	1-4 vs 5- 7	0.9	1.2	0.7
	5-7 vs 8-10	0.2	0.7	-
	1-4 vs 8-10	1.1	1.2	0.9
I	1-4 vs 5- 7	2.2	2.2	1.0
	5-7 vs 8-10	1.0	2.1	0.5
	1-4 vs 8-10	1.2	2.4	0.5
e	1-4 vs 5- 7	0.4	0.3	1.3
	5-7 vs 8-10	0.0	-	-
	1-4 vs 8-10	0.4	0.3	1.3
ɛ	1-4 vs 5- 7	3.5	5.2	0.7
	5-7 vs 8-10	4.1	4.6	0.9
	1-4 vs 8-10	0.6	5.1	0.1
æ	1-4 vs 5- 7	5.3	7.8	0.7
	5-7 vs 8-10	2.5	6.8	0.4
	1-4 vs 8-10	2.8	7.6	0.4
ʌ	1-4 vs 5- 7	2.1	3.8	0.6
	5-7 vs 8-10	2.1	3.7	0.6
	1-4 vs 8-10	0.0	-	-
ɔ and u	1-4 vs 5- 7	0.2	0.2	1.0
	5-7 vs 8-10	0.0	-	-
	1-4 vs 8-10	0.2	0.2	1.0
ʊ	1-4 vs 5- 7	0.6	1.0	0.6
	5-7 vs 8-10	0.3	0.3	1.0
	1-4 vs 8-10	0.9	0.9	1.0

In some phonetic classifications middle vowels are grouped with back vowels. On this basis, 92 per cent of all vowels uttered by babies are front vowels and 8 per cent are back vowels. The ratio is about 9 to 1. On this basis also adults use about 49 per cent front vowels and 51 per cent back vowels, a ratio of 1 to 1. These ratios, however, should be used with caution. They are more suggestive than conclusive, for more comparable values would be furnished by a comparison of the non-crying sounds of newborns with those of the adult group. These data, however, are not yet available.

Regardless of what phonetic classification is used there still is

good indication of the probable course of vowel usage during infancy and childhood. Apparently the crucial aspect of the change lies in the use of back vowels. The actual stages of this development are now under investigation.

SUMMARY AND CONCLUSIONS

Forty infants, twenty boys and twenty girls, under ten days of age were used for the purpose of studying the vowel sound equipment of newborns. Vocalization was entirely of a crying nature. The vowel elements appearing on each of twenty-five respirations were transcribed in the symbols of the International Phonetic Alphabet simultaneously and independently by two observers. A total of 1,000 respirations constitutes the data of the study. It was found that:

1. The time sampling method is inadequate.
2. A behavior unit method (the single respiration) of observation enables the observer to easily transcribe such sounds as occur during the unit.
3. Agreement between observers averages 85 per cent and in certain situations is well over 90 per cent.
4. Only four vowels occur with appreciable frequency in newborns: I, ɛ, ɜ̞, and ʌ.
5. The most frequently used vowel element in the vocalization of newborns is the sound ɜ̞. It is the only vowel sound formed by all forty babies.
6. Front vowels are more frequently used than middle and back vowels.
7. Among the front vowels the i and e sounds are least used.
8. Back vowels are almost entirely absent.
9. There are no sex differences in the use of vowels.
10. There are no age differences within the period studied.
11. A comparison of these data with data in vowel usage in adult speech reveals striking differences.

It is inferred from these data that the course of development during infancy in the use of vowel sounds is the increasing use of back vowels.

REFERENCES

(1) Fairbanks, Grant: Voice and articulation drillbook, New York, Harpers, 1940, pp. xxiv, 234.

(2) Irwin, Orvis C.: Speech sounds of infants during the first six months of life. Accepted for publication in the Psychological Bulletin.

(3) Irwin, Orvis C., and Krehbiel, Thelma E.: Speech sounds of infants: The fourth, fifth, and sixth months of life. Unpublished study.

(4) Voelker, Charles H.: A comparative study of investigations of phonetic dispersion in connected American speech. Extract des Archives Neerlandaises de Phonétique Experimentale, 1937, 13, 138-152.

THE PROFILE AS A VISUAL DEVICE FOR INDICATING CENTRAL TENDENCIES IN SPEECH DATA

ORVIS C. IRWIN[1]

The purpose of this article is to describe a graphic method for simultaneous presentation of central tendencies in data containing a large number of categories such as are characteristic of speech sounds. It is not a new device. Rather it represents an application to transcriptions of vowel and consonant sounds of a device successfully used with other kinds of data. The method is especially useful for comparing individual and group differences in the production of the speech elements and for following trends of development. It consists of a profile of the relative frequency of speech elements or categories of elements and involves an x and a y axis. However, it is a visual aid rather than a system of co-ordinates. Its construction consists of a simple modification of the principle of the bar graph. A series of equidistant points is laid off on the x axis from which perpendiculars may be erected to accommodate bars which signify the frequencies of the various elements. The y axis is taken to represent per cents of frequency. The units of both axes are arbitrarily chosen, but for comparative purposes they should be kept equal from profile to profile. No assumptions concerning discreteness or continuity need be made about the units plotted on the x axis.

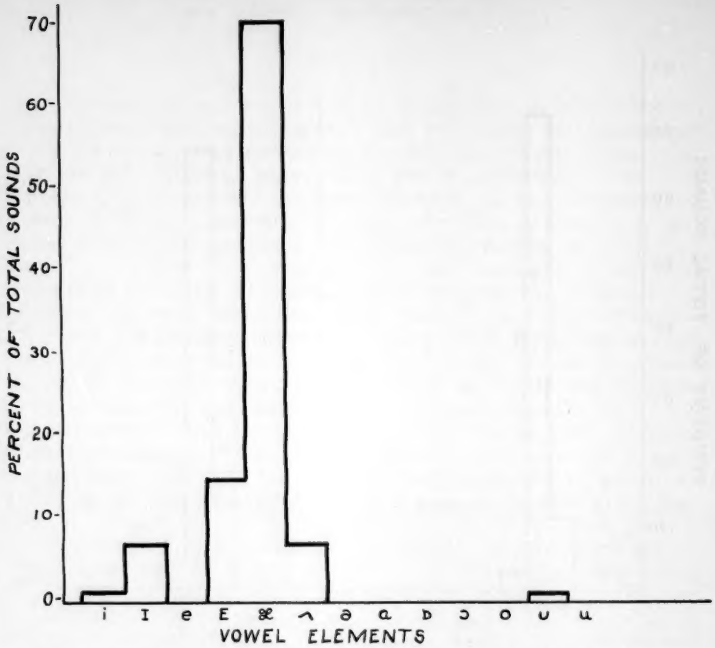
The essential condition for the usefulness of this device is the establishment of equal total areas under each of the compared profiles. The total area is taken as 100 per cent and represents the total frequencies of all sounds in each of the compared sets of elements. Thus the frequency of each speech element under the profile is in turn an area which represents its proportionate part of 100 per cent. Under these conditions the profile pattern will reveal quite readily the overall similarities and differences existing among groups of data.

This article will be devoted to various simple illustrative uses of the method. The profiles presented here are based on data which have been transcribed from the vocalization of newborn infants, of infants at several ages up to two years, of a group of low-grade feeble-minded children under five years, and from a study (2) on the distribution of elements in adult speech. A few of these samples are small, but since the aim is to illustrate a method and not to validate conclusions, the smallness of the sample does not prejudice the purpose of the article. The research findings as such may be found in reports appearing elsewhere. Moreover, it should be kept in mind that the profile illustrates only central tendencies and none of the other statistical characteristics of the data.

An analysis (1)[2] was made of the vowel elements occurring in the crying vocalization of forty infants under ten days of age. Figure 1 gives the data and illustrates the manner in which a profile is constructed. Thirteen vowel sounds are equally spaced on the x axis. The

[1] *From Iowa Child Welfare Research Station, University of Iowa, Iowa City, Iowa.*

[2] *A new method of collecting speech data involves a variation of the short sample technique consisting in a behavior unit rather than a time unit.*



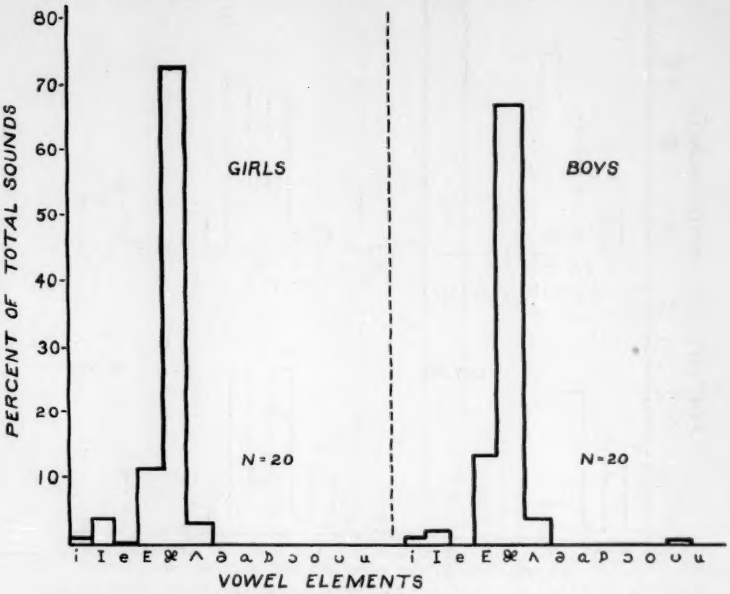
PROFILE OF VOWEL ELEMENTS DURING CRYING FOR FORTY NEWBORN INFANTS

FIGURE 1

y axis indicates the per cent of total vowel sounds. The area under the profile is taken as unity or 100 per cent. It will be noticed that in this profile 70 per cent of its area is occupied by the vowel æ. Fifteen per cent of the area represents the frequency of production of the vowel ɛ, and areas equal to 6 per cent each account for the vowels ɪ and ʌ.

This profile illustrates several interesting facts about the frequency of occurrence of vowel sounds in the crying of newborns: (1) that only four vowels occur with any appreciable frequency -- ɪ, ɛ, æ, and ʌ; (2) that the sound æ overwhelmingly dominates all other vowel sounds; and (3) that back vowels are practically absent from the sound repertoire of these forty babies.

The profile may be used to compare similarities and differences in certain aspects of two or more sets of data. Figure 2 presents two profiles, one the vowel equipment of girls under ten days and the other of boys. The patterns of distribution under these profiles are quite similar. They indicate that the outstanding predominance of one of the



PROFILES OF VOWEL ELEMENTS FOR GIRLS AND BOYS DURING THE FIRST TEN DAYS OF LIFE

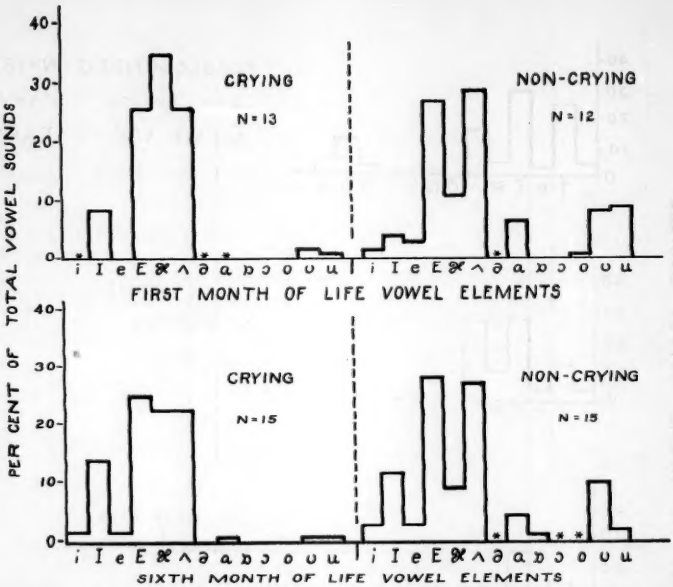
FIGURE 2

front vowels, as well as the absence of back vowels are characteristic of the crying vocalization of both newborn boys and girls.

The profile lends itself as a device for revealing age differences. Figure 3 shows four profiles illustrating differences between crying and non-crying vowel production of infants at one and six months of age.

Before considering this group of profiles, perhaps a preliminary comparison of the profile in Figure 1 should be made with the upper left profile in Figure 3. The former, as was pointed out above, is the dispersion of vowels of crying newborns, the latter represents the distribution of crying sounds of infants during the last half of the first month. Inspection of the two profiles reveals that whereas in the first the largest area, that representing æ, includes 70 per cent of the total area, this value in the second profile has been reduced to 35 per cent, or to one half, and that the ɛ and ʌ sounds are replacing it. Moreover, among the crying sounds of month-old babies there still are few back vowels produced.

Comparison of the profiles of first and sixth month crying infants



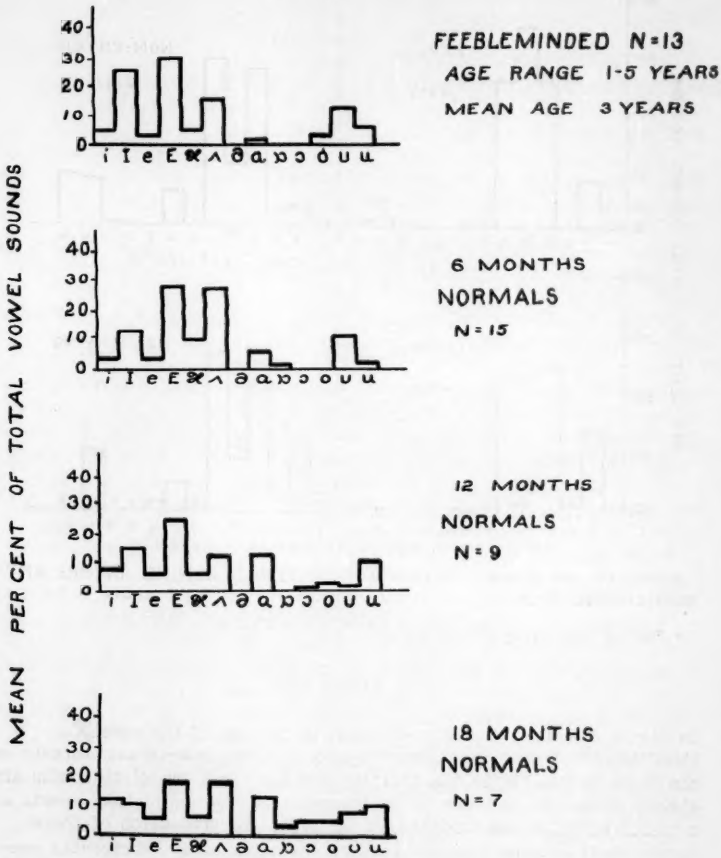
PROFILES OF VOWEL ELEMENTS UTTERED DURING CRYING AND NON-CRYING PERIODS FOR THE FIRST AND SIXTH MONTH OF LIFE

* *Per cent too negligible to indicate*

FIGURE 3

in Figure 3 shows a further reduction in the use of the vowel æ. Otherwise the two patterns are similar. In the non-crying infants on the right in the figure the profiles for the first and sixth months are almost alike. While there is an increment in the use of back vowels as compared with the two profiles at the left, the production of these sounds still remains a meager part of the total area. Horizontal comparisons of the profiles show that the vowel æ is more frequently used in the first month by crying than by non-crying children. Comparison of the two profiles for the sixth month shows a similar situation.

Figure 4 gives another age distribution of vowels. The three lower profiles present data of vowel sounds of normal babies at six, twelve, and eighteen months. An examination of the profiles for these age groups indicates a tendency toward an increase with age of the production of back vowels. It is manifested by a tendency toward equalization of the unit areas on the x axis. Thus the lowest profile in the figure is flatter than those above it. Some indication of the fattening process is beginning to appear in the profile of the twelve month group. It is quite appreciable in the eighteen month group.



PROFILE COMPARISONS OF VOWEL PRODUCTION OF FEEBLEMINDED AND NORMAL CHILDREN

FIGURE 4

An exceedingly interesting comparison is afforded by the first profile in the figure with each of the other three. Consider first the contrast between the upper and lower profiles. The upper one represents the vowel usage of a group of low-grade feeble-minded children varying in age from one to five years with an average age of three

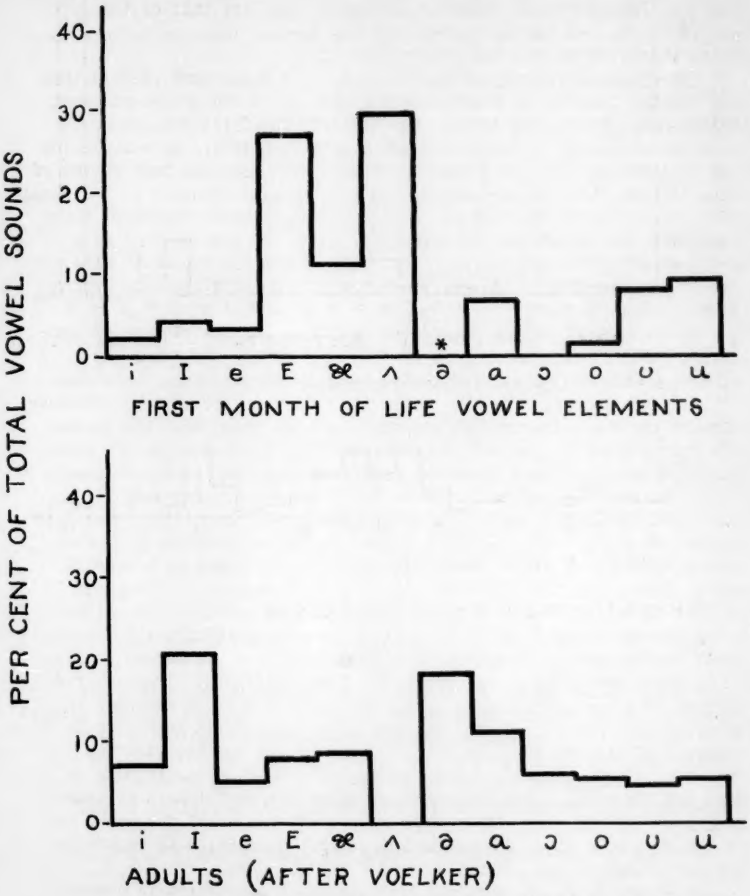
years. The lower profile is that of the one and one-half year old group of normal babies. The greater part of the area of the profile for the feeble-minded group is located in the left half of the profile. The area in the profile for the normals tends to be more evenly distributed over the entire vowel range.

The upper profile resembles the second and third more closely than it does the fourth. It approaches the pattern of the second somewhat better than it does the third. The areas of the first and second profiles both show preponderance in their left halves. By calculating the percentages from the y axes it will be seen that the left halves of both include about 80 per cent of the total areas. Thus it is indicated that the central tendencies in the production of vowel sounds of these low-grade feeble-minded children is about the same as that of a group of normal babies in the second half of the first year of life. The comparative use of the profile renders this quite apparent from a visual standpoint.

Figure 5 presents two dissimilar profiles, one the picture of non-crying vowel production of infants under one month, the other the pattern of adult speech. The latter is an adaptation from vowel data collected by Voelker (2). The following are the outstanding differences between the two patterns: (1) the upper profile shows that the sounds are concentrated in the left half whereas in the second they are more evenly distributed over the vowel range; (2) the most prominent vowels in the first are ɛ and ʌ, in the second I is the most frequent; (3) more back vowels are present in adult speech; (4) the middle vowel ʌ is present in infant vocalization but ə is absent, while in adult speech this dispersion of middle vowels is reversed; (5) there is a bimodal distribution among the front vowels of infants whereas in the adult profile there is revealed a tendency among front vowels to be unimodal.

A further comparison of non-crying sounds of infants with adult vowel production may be made from Figures 4 and 5. The second profile in Figure 4 shows the vowel usage of six month infants. It is characterized by a concentration of vowels in the left half in contrast to the flatter distribution of adult vowels in the profile in Figure 5. In general, of all the profiles in Figure 4, the one for the eighteen month distribution presents the closest approximation in flatness to that for the adult. The comparison of these two profiles is at least suggestive of the possibility that adult status of vowel frequency may be achieved soon after the eighteenth month. A decision in regard to the matter must await the collection of further data.

Irwin and Curry (1) observed that consonant production in newborn infants is quite infrequent. Consequently a profile of these speech elements is not yet available for this age. However there are some data available for the frequency of consonant usage by a small group of six month infants who were not crying. Figure 6 presents a profile constructed from these data. Below it is a profile of adult consonant production adapted from the Voelker data. The two patterns are hardly alike. Of the twenty-three consonant sounds only twelve, or about half, are produced by six month infants, whereas all are used by adults. An interesting contrast exists in the frequency of the consonant h. In infants this sound occupies 64 per cent or about two-thirds of the



PER CENT FREQUENCY OF VOWEL SOUNDS

* Per cent too negligible to indicate

FIGURE 5

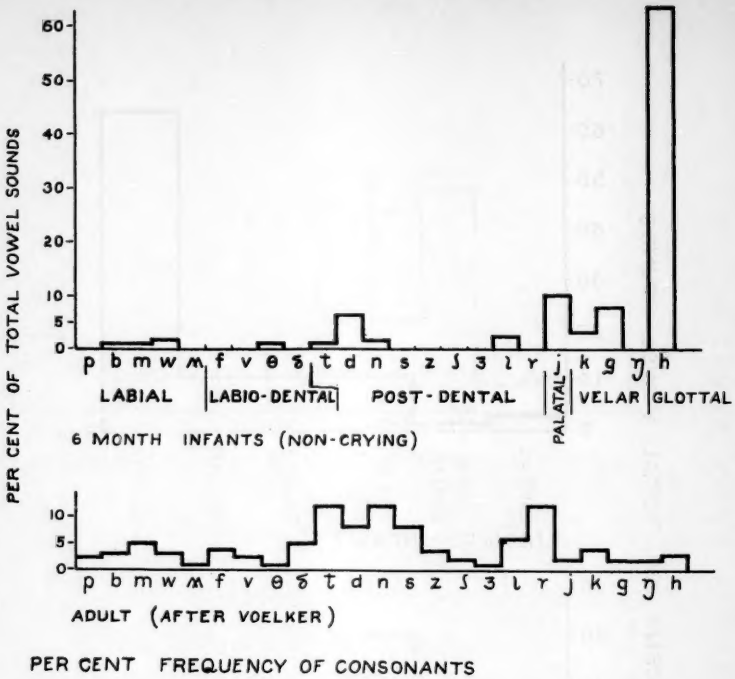


FIGURE 6

area under the profile. In adults the corresponding area is 3 per cent. The sound h, of course, requires little mouth formation, but the difference between its frequency in infants and adults is exhibited in the profile in a striking manner. If this sound were eliminated from the profile on the ground that there is little in the way of special mouth formation, the consonant profile still would not be similar to that of the adult.

In Figure 7 the consonant elements are grouped into classes according to the place of articulation. The upper profile indicates that infants produce consonants somewhat more frequently with the rear parts of the mouth than do adults. The lower profile shows the greater extent to which dental consonants are produced by adults than by infants, and that palatal, velar, and glottal structures are less frequently employed in adult consonant output.

This is not the place to elaborate a possible interpretation of these particular profiles, but if they are at all representative it is not unlikely that any interpretation of them concerning the course of development must recognize the infrequent production by infants of both back vowels and of labial and dental consonants. A comparison of the infant status of these elements with their adult status may indicate a

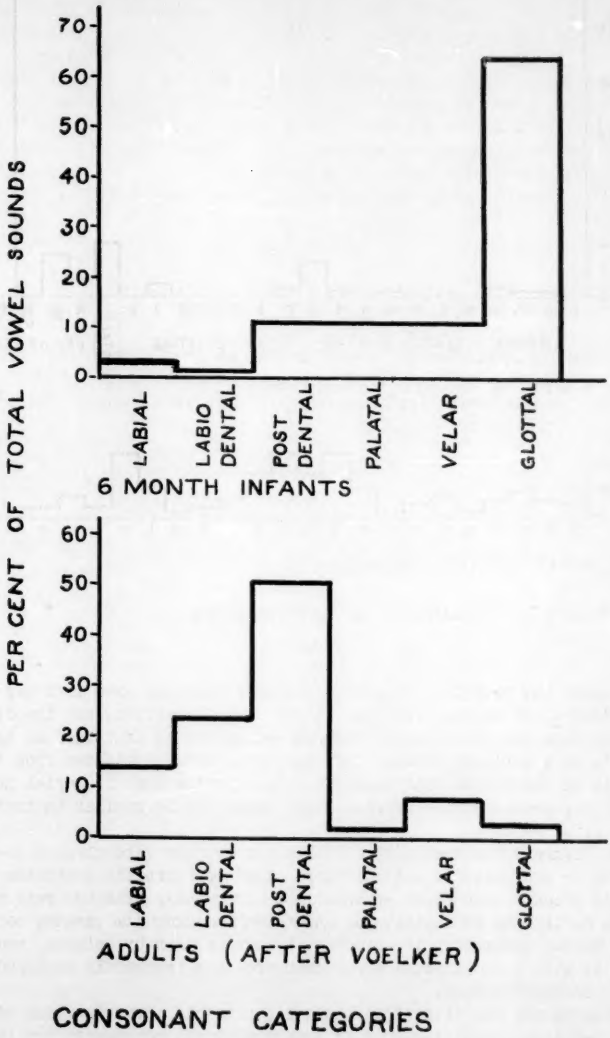

PER CENT OF TOTAL VOWEL SOUNDS
70
60
50
40
30
20
10
0
LABIAL
LABIO DENTAL
POST DENTAL
PALATAL
VELAR
GLOTTAL
6 MONTH INFANTS
60
50
40
30
20
10
0
LABIAL
LABIO DENTAL
POST DENTAL
PALATAL
VELAR
GLOTTAL
ADULTS (AFTER VOELKER)
CONSONANT CATEGORIES

FIGURE 7

possible future course of development of the newborn's speech sounds.

The object of this paper rather has been to illustrate the use of the profile with a type of data which is subdivided into a number of categories. Its advantage lies especially in its ability to present the overall effect of such data so that this effect is not lost in a detailed analysis of the component categories. The analysis of the development of speech sounds requires a device for preserving the general overview. The profile therefore has much to commend it as a graphic presentation of the central tendencies of these data, but it is not without its limitations, the greatest of which is the fact that it does not indicate variability within the individual category.

REFERENCES

(1) Irwin, Orvis C., and Curry, Thayer: Vowel elements in the crying vocalization of infants under ten days of age. Child Develop., 1941, 12, 99-109.

(2) Voelker, Charles H.: A comparative study of investigations of phonetic dispersion in connected American speech. Extract des Archives Neerlandaises de Phonétiques Expérimentale, 1937, 13, 138-152.

THE INFLUENCE OF SOCIAL FACTORS UPON THE PERSISTENCE OF SATIATION IN PRE-SCHOOL CHILDREN[1]

ARTHUR BURTON

One of the factors in industry lowering efficiency and reducing output is satiation or monotony.[2] Although satiation is principally fostered by activities of a repetitive and uniform nature, the conditions under which the work is carried out play a considerable part in the general reaction to monotonous work. One of the most important of these work conditions is the relation of face-to-face employees with each other. For example, Wyatt, Frost, and Stock (7), in a study for the Industrial Health Research Board, found a high degree of similarity in the production curves of employees working together; a fact which they attributed to the suggestive influence of the group. The group influence on production was remarkably demonstrated when selected operatives were paired for pre-determined work periods, and then isolated for similar periods. Personal characteristics, e.g. intelligence, temperament, etc., were causal elements in the facilitating or depressing effect on production. Davies (2) reports an instance where boredom in a small group of operatives seems to have spread to the entire group of employees producing general dissatisfaction with the work. Taylor, Thompson and Spassoff (5) report no facilitation in performance or reduction in boredom or tiredness when subjects working alone were compared with those working in groups. If industry is interested in maintaining efficiency, it must provide work conditions which will delay the onset and prevent the persistence of satiation. This study, while not an industrial one, concerns itself with one such factor.

The determination of the effect of social rather than individual work situations on satiation has more than industrial significance. Satiation is a common behavioral phenomenon which usually, but not necessarily, appears when the motivation governing an act is diminished or entirely lost. If the factors governing the onset and persistence of satiation be uncovered, then an important contribution can be made to the psychology of motivation.

The present problem was first motivated by the observation of an absence of satiation in the general activity of pre-school children on the playground. This may be due to the multiplicity of social contacts and the abundance of play material. While both variables probably contribute to the absence of satiation effect, the former seemed to us of greater importance. An experiment was therefore designed with the objective of determining the value of social stimulation in the reduction of satiation in pre-school children. An indirect objective was the desire to elaborate a technique which could be used at a number of age levels.

[1] *From the Institute of Child Welfare, University of California.*

The author gratefully acknowledges the suggestions and criticisms of Professor Harold E. Jones, Director of the Institute of child Welfare. Acknowledgment is also made to the Works Progress Administration for assistance in recording certain observations.

[2] *A variety of terms have from time to time been employed to represent the satiation process. Monotony and boredom are favored in industry, but satiation is more common in laboratory studies or those influenced by Lewin. The latter will be employed here with certain exceptions when reference is made to industrial studies.*

EXPERIMENTAL DESIGN

The rationale of the experimental design was as follows: The subject (S_1) was satiated upon a simple, uniform and repetitive task. The criterion of satiation - as in previous studies - was the rejection and non-resumption (for at least 10 minutes) of the task when this rejection was not due to extraneous needs i.e., toilet, sleep, hunger, etc. needs. Following satiation, and the complete unwillingness of S to continue the task, a second motivated child (S_2) was brought in to help S_1 finish his "game." The order of magnitude of the activity resumed by the satiated child (S_1), under the social influence of S_2, was computed as indicating the degree of freedom from the persistence of satiation on the identical task.

The subjects for this experiment were pre-school children attending the nursery school (1938-39) attached to the University of California Institute of Child Welfare. Twenty-four children - 10 males and 14 females - were available from the personnel of the morning and afternoon sessions after non-participants were eliminated for reasons of immaturity, emotionality, and schedule difficulties. The mean chronological age was 45.3 months, and the mean mental age 54.8 months. The range of ages was 31 to 56.5 months. On the whole, the children were from a superior socio-economic environment.

For this experiment, a satiation peg-board was devised. It consisted of a flat, rectangular board composed of three sections, each 22 by 11 3/4 inches (See Figure 1). The board was hinged so that any or all of the sections might be exposed to view at one time. Each section contained 325 pockets constructed to house a standard Bradley and Co. 1/8 inch peg. The surface of each pocket was outlined in white for greater visibility. The range of insertions possible for each subject was 1 to 975. In actual practice only one child succeeded in filling all three sections. As a rule, two sections of the board constantly were exposed; if the S completed the two sections, the third was opened. One thousand pegs in four primary colors accompanied the satiation peg-board. The pegs were in a small box at either side of the board, contingent upon the hand preference involved. When 2 S's were at work, the pegs were evenly divided on the right and left hand side (Figure 1). A Bristol Recorder was used to record the number and rate of insertions; the number of 30-second stoppages voluntarily adopted was later calculated from this record. A W.P.A. assistant aided in securing protocols when E was occupied.

The experimenter (E) selected a child from the nursery-yard who was sufficiently mature physically for experimental purposes, emotionally stable,[3] was not sleepy, and above all, showed a minimum of motivation in regard to any activity or toy in the nursery school.[4]

[3] *Preliminary trials had confirmed our expectations that the involved nature of the experiment would provide some emotional strain. In fact, as will be seen later, the emotional rejection of the task through anger, crying, or leaving the field constitutes one criterion of satiation. On the recommendation of the nursery school supervisor certain "emotional" children were excused from the experiment.*

[4] *In experiments where the measurement of motives is a requisite, it is desirable to locate a basal point in the form of a zero or neutral state of motivation. Without such a zero point absolute magnitudes and scalar units lose universal meaning. The attainment of such a neutral state implies the measurement and quantification of all instigatory behavior in the individual at any given moment. Unfortunately, this remains a problem for psychology of the future. Under the circumstances, the most reasonable procedure open to us was to select subjects unoccupied at the moment and then attempt to motivate each child equally.*

A

B

C

D

E

F

Figure 1

S_1 was asked if he would like to play a "new" game. If the child agreed, and appeared interested, he was taken to the toilet and then to the experimental room. There he found the peg-board with the pegs. S_1 was told that this was a "new" game in which all the pegs had to be put in the board. E illustrated by inserting two pegs; and the child was encouraged to try it. After S had inserted a few pegs, he was told that E had filled the entire board with pegs but somebody "mean" had come and taken them all out. "Would he help E put them in?" The response to this plea was usually indignation and a willing acceptance of the job. E stayed in the experimental room rarely more than four minutes; only sufficient time was taken to insure good rapport and to offer a word of encouragement once activity was under way. At the end of this short period, E told S that he had to leave the room to speak to the teacher, but that he would be back soon. S could keep right on playing and put all the pegs in, but he was not to leave the room himself. E then stepped into the adjoining room and observed S through a one-way vision screen. Only one subject (of a total of 25) failed to respond to these instructions, and had to be excused.

E returned to the experimental room only when the criterion of satiation was attained. The criterion, as previously mentioned, was the rejection of the task and non-resumption for at least a period of 10 minutes. If the child left the field, i. e., opened the door and went downstairs contrary to E's orders, this was considered as satisfying the criterion since the pressure of satiation was sufficient to overcome the rigid code of compliance. The validity of the criterion is partially established by the observations that once the child left the field, it was difficult to induce him to come back. Those remaining, following cessation of activity, would verbalize their dislike of the task, frequently reiterating that they were "tired," and presenting a typical picture of a satiated individual. The adoption of innumerable irrelevant goals, i. e., creative substitution, is further evidence of satiation.

When E returned to the experimental room, the S in a satiated state, the assistant left her post behind the oneway vision screen to obtain the second child (S_2) who was to act as the social influence. S_2, the second child, was selected on the same basis as S_1. Because of the rigid requirements of the experiment, and because of certain limitations imposed by the nursery school schedule, it was impossible to pair the subjects in any rigidly consistent way, i. e., having a single child act as a constant social influence. However, no child served as an S_2 before he was an S_1, thus eliminating the effects of re-satiation (übersättigung)(4). It was intended that no child should serve more than once as S_2, but in a few instances exceptions had to be made to this rule. These exceptions seem to be of little consequence since S_2 was always well motivated and "fresh" from the playground; in addition, only a 15-minute social period was required which is usually not sufficient to re-satiate a motivated individual.

While S_2 was being brought to the experimental room, E maintained rapport with S_1, at the same time recording the pegs which had been inserted. S_1 was told that someone was coming to play with him. E left the room, returning to the observation screen. The assistant

entered with S_2 and repeated the original instructions with the addition that S_2 would now help him finish the "game." The assistant left and returned to the screen. A 15-minute time limit was established for this part of the experiment.

Figure 1 is a series of individual frames of a 16 mm. film illustrating the performance of a single subject. In Frame A activity has proceeded about 3 minutes. Frame B represents about 7 minutes. Satiation is beginning to set in (Frame C) and S_1 is "blubbering". S_2 has been introduced in Frame D, and they continue to fill the peg-board. One section has been completed and a new one presented (Frame E). S_1 is again motivated. Aggression appears terminally in Frame F.

COMPARATIVE RESULTS

The more qualitative aspects of this study are reserved for discussion elsewhere (1).

In this and the following sections, we shall speak of the initial period of activity--in which S_1 was satiated--as the <u>satiation period</u>; the interval in which the social stimulation was applied by S_2 will be known as the <u>post-satiation period</u>. It should be understood that although the initial period is known as the satiation period, it involves a transition from motivation or pre-satiation to satiation.

The average number of pegs placed prior to satiation was 310.1 $\pm$ 35.0; the P. E. of the distribution was 167.8 and the range 54-975 (Table 1). The mean satiation time was 34 min. 46 sec. $\pm$ 2 min. 14 sec.; the P. E. of the distribution, 10 min. 44 sec.; with a range of 9 min. 1 sec. to 60 min. 40 sec. The mean number of 30 second pauses voluntarily adopted by the subject during the satiation period was 13.8 $\pm$ 1.6; the P. E. of the distribution 7.2 and the range 1-46. This may be interpreted as indicating that a mean period of 7 min. or 20 per cent of the mean time to become satiated was spent in some form of distraction.

The number of pegs inserted and the pauses voluntarily adopted during the 15 min. post-satiation period is the quantum index of the effect of social stimulation on satiation. In the post-satiation period the mean number of pegs inserted was 45.4 $\pm$ 4.7, the P. E. of the distribution was 23.1 and the range 4-104. The mean number of post-satiation 30 sec. pauses was 19. $\pm$ 1.1; the P. E. of the distribution 5., and the range 5-33. Sixty-three per cent of the total average satiation time was spent in inactivity.

The satiation and the post-satiation periods cannot be directly compared because of the inequality in mean performance time. However, by comparing the post-satiation period with the initial 15 minutes of the satiation period, some index may be obtained of the resumption value of a social influence. Since the subject may be expected to be most highly motivated during the initial period of motivation to satiation (and ipso facto, least satiated), the degree to which the performance quanta under the social influence approaches the initial period of putatively highest motivation thus becomes the index.

The mean number of insertions for the initial 15 min. of the satiation period was 139.1 $\pm$ 7.3, and 48. $\pm$ 5.5 for the coordinate post-satiation period. The resumption value of the post-satiation period is

TABLE 1

SUMMARY OF INSERTIONS, DURATION, AND PAUSES REQUIRED TO SATIATE AND TO STIMULATE RESUMPTION OF ACTIVITY FOLLOWING SATIATION.*

Subject	No. of Insertions to Satiate	Satiation Time	Pauses (30 Secs.)	Post-satiation Social Stimulus	Post-satiation Insertions (15 Mins.)	Post-satiation Pauses (30 Secs.)
B.E.	324	45'27"	19	S.Hu.	104	6
H.C.	54	9' 1"	1	B.E.	20	23
A.G.	156	22'52"	10	S.M.	46	18
S.M.	137	23'49"	15	M.E.	10	27
S.H.	173	31'13"	20	S.Hu.	26	21
J.E.W.	86	20'11"	24	M.A.	4	27
D.S.	144	17'57"	3	S.B.	9	23
S.B.	489	55'17"	7	A.C.	14	33
E.V.	350	39'38"	15	A.G.	75	16
S.Hu.	68	9'44"	8	A.Ho.	6	27
H.C.	975	58'47"	4	E.V.	9	25
B.B.	324	34'15"	9	D.S.	77	17
G.R.	848	56' 2"	5	D.D.	79	13
M.E.	325	31'52"	5	B.E.	55	16
M.K.	192	31'54"	24	M.P.	20	23
M.P.	122	23' 4"	19	B.B.	12	20
M.R.	503	47'50"	11	M.K.	68	14
J.T.	94	41'12"	46	L.F.	29	17
B.C.	149	34'00"	18	G.R.	4	27
A.Ho.	650	60'40"	3	M.W.	67	5
S.A.	207	26'32"	--	J.D.	100	--
J.M.	323	37'25"	31	S.A.	87	24
R.H.	100	11'00"	--	M.K.	76	6
M.W.	650	56'42"	7	L.F.	92	10
Mean	310.12	34'46"	13.82	Mean	45.38	19.04
$P.E._D$	167.80	10'44"	7.15	$P.E._D$	23.14	4.98
$P.E._M$	34.96	2'14"	1.56	$P.E._M$	4.72	1.06

A standard time limit of 15 minutes was adopted for the post-satiation period.

34 per cent when the mean of the initial 15 min. satiation period is considered the highest point of motivation. Likewise, the mean number of 30-sec. pauses initially was 3.9 ± .63 and finally 19. ± 1.1. The amount of time lost in stoppages in the post-satiation period was 48.7 per cent of that lost in the 15 min. initial satiation period. The differences between means are statistically reliable by conventional formulae (3).

The differences between the satiation and post-satiation groups is best shown by means of a graph (Figure 2). The mean satiation curve for the pre-satiation period shows a steady decrease in the number of insertions as satiation is approached. The individual work curves indicate the high degree of variability. The differences between the means of the plotted points for the initial and final 15-min. periods are statistically significant.

A comparison was made of the rate of work in the two periods. These data are plotted in Figure 3. Here again individual differences are great. The rate is initially 9.17 insertions per minute; then drops, followed by a rise exceeding the opening rate. After 9 minutes of activity, there is a steady decline in rate of work with almost a plateau from 21 to 30 minutes. These curves differ slightly from those

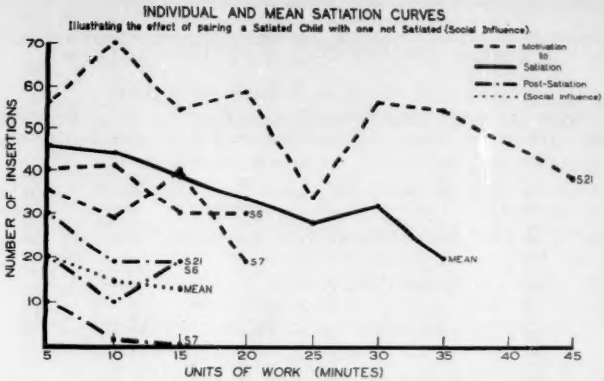


Figure 2

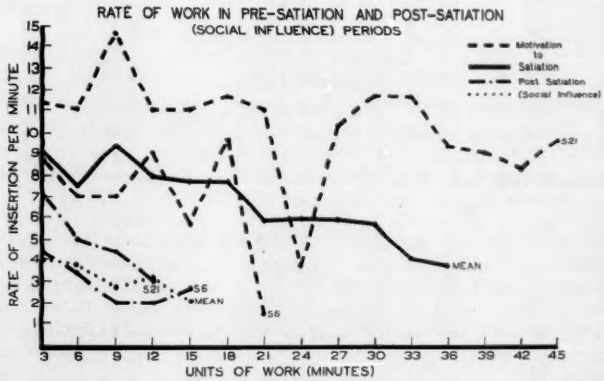


Figure 3

reported for industrial workers (6-8). There is no mid-spell depression evident for our S's worked only to the point where they individually became satiated and not for long arbitrarily set periods.

Inspection of the individual curves, and the qualitative observation, points out that as a satiated state is approached many subjects show great alternation in rate - they work in spurts. Others succeed in maintaining a fairly constant rate up to the very point of rejection; still others offer a constant rate of work, but pause frequently and for relatively long durations.

The rate of peg insertion in the social period is slower than when working alone and decreases with further effort.

DISCUSSION OF RESULTS

The validity of the state of satiation induced in our subjects is insured by its conformity to available criteria. In addition to rejection and non-resumption of the task (4), the qualitative indices - postural, facial, and verbal - indicate satiation.

When the mean number of insertions for the initial 15 minutes of the satiation period - the point where motivation was presumably greatest - is considered the highest obtainable resumption value (100 per cent) on the insertion scale, a mean resumption value of 34 per cent was obtained for the post-satiation period. Assumed at face value, we may say that one-third of the original instigation directed toward the solution of the task was revived with social stimulation. Stated still another way: one-third of the accumulated effects of satiation was removed by such treatment.

Deducting the time spent in distractions and considering the mean number of minutes spent on the initial 15 minutes of the satiation period as 100 per cent resumption, a resumption value of 43 per cent was obtained for the post-satiation period. In terms of average time spent in activity, more than two-fifths of the original average instigation was recovered, or the satiation effects dissipated.

In the attempt to furnish this study with quantitative precision we have given little attention to the individual differences involved. In the sense of central tendency, the preceding data indicate a substantial loss of satiation following subjection to social stimulation. But sober observation and inspection of individual results reveal that some children recover rapidly from their satiation in the presence of another child; others recover only slightly; and still others, not at all. In a given group, albeit the general tendency is for resumption, any single child may or may not react to the social stimulation. The determination of the causal variables which predispose to a reduction in satiation with social stimulation will be left for consideration in a later study.

Consideration of the findings of this study must, in addition, take account of the small sample involved and the limited nature of the social pairings.

Despite differences in criteria of satiation, in an industrial versus laboratory setting, in adult salaried employees rather than nursery school children, and the principal difference of pairing workers before satiation rather than after, the findings of Wyatt, Frost, and Stock (8) are in harmony with the present study. They concluded from their intensive factory study: "In the present inquiry the arrangement of workers was similar to the usual factory procedure and it was obvious that the rate of working of each individual was closely dependent on that of her neighbor and upon the behavior of the group as a whole." (p. 51)

In conclusion, it may be said that the involvement of fellow individuals in the immediate work or play situation acts, on the whole, in the direction of relieving the persistence of satiation. Individual

differences and variability are grossly apparent and necessarily qualify the generalization.

SUMMARY

1. Twenty-four nursery school children were satiated upon a repetitive and uniform task - peg insertions - to determine whether social stimulation would reduce the persistence of satiation.

2. The mean time to become satiated was 34 min. 4 sec. ± 2 min. 14 sec. The mean number of insertions was 310.12 ± 34.96.

3. The resumption value after satiation for peg insertions was 34 per cent of the number inserted during an equivalent period of modal motivation.

4. The resumption value in terms of time spent in activity, exclusive of distractions, was 43 per cent of an equivalent period of highest motivation.

5. It was concluded that the social stimulation afforded by an associate after satiation acts in the direction of reducing the persistence of satiation. Individual differences are great and there are many exceptions to the rule.

6. The results of this study are, on the whole, in harmony with the findings of Wyatt, Frost, and Stock on industrial workers.

7. The hypothesis was substantiated that social factors tend to account for the general absence of satiation in the nursery school.

REFERENCES

(1) Burton, A.: The aggression of pre-school children following satiation (in print).

(2) Davies, A. H.: The physical and mental effects of monotony in modern industry. Brit. Med. J., 1926, 2, 472-479.

(3) Garrett, H. E.: Statistics in Psychology and Education, (2nd Edition). New York, Longmans, Green Co., 1939.

(4) Karsten, A.: Untersuchungen zur Handlungs- und Affekt Psychologie: V., Psychische Sättigung. Psychol. Forsch., 1928, 10, 142-254.

(5) Taylor, J. H., Thompson, C. E., and Spassoff, D.: The effect of conditions of work and various suggested attitudes on production and reported feelings of tiredness and boredness. J. Appl. Psychol., 1937, 21, 431-450.

(6) Wyatt, S.: Boredom in industry. Personnel J., 1929, 8, 161-171.

(7) Wyatt, S., Frost, L., and Stock, F. G. L.: Incentives in repetitive work. Ind. Heal. Res. Bd., Rep. No. 69, 1934.

THE MEASUREMENT OF PARENT BEHAVIOR[1]

HORACE CHAMPNEY

It is the purpose of this paper to describe the Fels Parent-Behavior Rating Scales. Whether the reader is interested in using the scales, in evaluating the results of their use in research, or merely in problems of measurement in social psychology, he will perhaps wish to know something of the rationale which lies behind the final product. With this in mind, the following topics are treated: the human judge as a psychometric instrument; the rating scale as a device for the expression of judgments, its strengths and weaknesses, and a reinterpretation of some classic objections; an improved rating-scale format; specifications for good "behavior cues"; and a procedure by which the cues may be scaled.

A. WHY A RATING SCALE?

Despite the subjectivity of the rating scale - and its alleged gross unreliability - we have adopted it as the basic technique for measuring parent behavior as observed by the home visitor. For various reasons, discussed below, we believe that the rating scale is not necessarily so unreliable, and that its very subjectivity may be a decided asset.

The Human Organism as a Precision Instrument

How accurately can a rater discriminate? In the field of simple sense perception it has long been known that the human organism is able, without the mediation of special instruments, to make rather fine distinctions. Brightness, pitch, temperature, and the two-point limen have been measured by the human sensorium, with the aid of the psychophysicist, in terms of quantitative scales of considerable refinement. Somewhat more recently it has been demonstrated that the number of discriminations of which the perceptual mechanism is capable may be substantially increased by experience or training. This is particularly true of the more complex perceptions, like tea tasting, perception of music, recognition of faces, discrimination of concepts through language, and the pre-laboratory type of medical diagnosis.

But somehow orthodox American psychology has largely refrained from capitalizing on the human organism as an instrument for precise measurement. The revolt of the behaviorists against introspection, and the accompanying drive to have psychology recognized as a biological science, has tended to imbue us with a distrust of any procedure which is even slightly tainted with the subjective. Interpretation, inference - and their dark cousin, Intuition - have been considered unscientific. Too often the job of the psychological observer has been limited to setting down in cold tally marks or stop-watch readings only such events as are

[1] *From the Samuel S. Fels Research Institute, Antioch College. This is the second of a series of papers on the development of the Fels Parent-Behavior Rating Scales (4). The first paper dealt with the selection of variables of parent behavior (7); the third will present data concerning norms, validity, and reliability based on three years' use of the scales in the Fels longitudinal study of 150 normal children (6). All three papers are, in part, an extension of the writer's doctoral dissertation presented to The Ohio State University under the advisorship of Professor Horace B. English (5).*

concretely obvious and unequivocal. We have been rigorously trained to preserve our higher perceptual powers for the service of strictly non-scientific philosophizing.[2]

In the field of child development some of the finest work of the past decade illustrates the objective approach; there are the studies of A. P. Weiss and his associates on infant behavior (19), the objective methods for studying social behavior in the nursery school developed under Dorothy Swaine Thomas at Columbia (23), and Myrtle McGraw's "condensed milk" technique for analyzing infant locomotion (16). Uniformly such studies have yielded measures characterized by high reliability; their significance for the larger problems of development and social psychology has been less clear.

When a variable has been so objectified that it can be tallied mechanically it may have been so throughly divested of its spatial and temporal context that the significant configuration is destroyed. The number of five-second periods out of thirty minutes of free play during which the child laughs - if the observers can learn to agree on what constitutes a laugh - might well yield a reliable score for frequency of laughter during the free play period in nursery school. To assume that it measures the child's sense of humor is quite another thing. It is conceivable that for certain sensory-motor problems the frequency of laughter would be a significant variable. But for a study of personality we may be more interested in a measure of sense of humor which is only fairly reliable than in a highly reliable measure of frequency of laughter.

Turning to the home environment, if we are interested in the severity of penalities employed in disciplining the child we might count the number of spanks administered over a stated interval of time under standard conditions. We might even rig up a device for measuring the number of foot-pounds of energy applied to the child's rear. These would be objective measures. But their value would be doubtful when divorced from the context of the immediate situation, the past history of similar situations, and the general tone of relationship between parent and child. A useful evaluation of any event, or series of events, in parental behavior calls for many allowances and corrections. How provocative was the behavior of the child? To what extent was the mother modifying her habitual mode of response because of the presence of the visitor? Many of the events which are crucial in revealing parent attitudes will seldom or never occur before the visitor's eyes. Such evidence is available only circumstantially, through inference from child reactions and casual conversation with the parent.

The integration and appraisal of so complex a mass of evidence is clearly outside the limitations of the objective approach. Human judgment, despite its liability to error, is the only instrument which can begin to meet the demands of the problem. Inference, far from being the fly in the ointment, must be cultivated for all it is worth. Much of it will be below the level of rational awareness, emerging as "intuitive" knowledge, unsupported by verbalizable observations. Just as the

[2] *The credo which enthrones Objective Measurement on the right hand of Science and sends Intuition to reign over the Lower Regions is at length becoming an historical curiosity. No more vigorous, entertaining, and provocative fulmination against this narrow view of the scientific method will be found than the recent book by David Lindsay Watson (25).*

seasoned old country doctor, quite without benefit of differential blood count or even thermometer, knows that the child has measles, so the experienced home visitor will know that the child is over-protected.

A well-written case report, presenting a thoroughly interpretive clinical picture of the child in his home environment, has much to be said in its favor, especially as part of an individual case-study approach. But, whereas it gains much in flexibility to fit the unique aspects of the individual child, it loses the advantages of quantitative standards and statistical analysis.

The rating scale would seem to be the happy compromise. It stays within the bounds of standardized variables and quantitative methods, yet it can give adequate scope for the expression of expert human judgment. It is perhaps the ideal device for evaluating those configurational qualities which do not readily reduce to the behavioristic atoms demanded by the objective method. The observer is released from his constraint to the mechanical tallying of the obvious, and challenged to employ all the perceptual and intuitive skill which comes of long training and experience.

The Unreliability of the Rating Scale

Some of the "unreliability" for which the rating scale technique is noted may be traced to the fact that ratings have been employed largely in attempting to measure the more difficult and elusive variables, those for which objective techniques have been impossible. There are two strikes on the pinch hitter when he goes to bat. But in addition to this there seems to have occurred a combination of historical circumstances leading to a degree of pessimism which is unwarranted. Rating scales are indeed being used, perhaps more than ever, and in some of the best research circles. But they are usually adopted with reluctance, as a last resort, and with the feeling that all hope for psychometric precision is thereby abandoned. Let us consider the sources of this point of view.

The accuracy with which a given rater can discriminate will certainly depend upon many things. Among these we should expect to find the rater's previous experience with the sort of material being rated, and his habits of attention to, and familiarity with, the concepts treated in the scales. Adequate pre-observation of the actual material he is to rate, with a deliberate focus on the appropriate variables, should also be important. Then there are doubtless individual differences between raters in their talent for making sound, integrated judgments uncolored by prejudice, preconception, and unwarranted inference.

Attention to these factors has, with few exceptions, been peculiarly weak in the history of rating-scale practice. Not that they have been overlooked; indeed they have been discussed at some length in textbooks on psychometrics (21, 14). But in actual service the majority of rating scales have depended upon raters without psychological or other comparable training in the scientific observation of human behavior. As a rule the rater has been concerned with other matters, the rating task occupying an incidental or even casual place in his routine. This is

clearly true of the army ratings, of most of the industrial ratings by foremen and supervisors, ratings obtained from employment references, and ratings by students and teachers.

When the post-war enthusiasm for rating scales had about reached a peak - a period when mass and casual ratings were particularly prevalent - two important papers appeared in the journals. An article by Max Freyd in 1923 (12) discussed in approving detail the merits of the graphic rating scale. This article probably remains the best treatment of the subject in the literature. About a year later P. M. Symonds published an article (22), based partly on a survey of rating scale experience and partly on certain theoretical derivations, concluding that the graphic scale achieved only illusory refinement, and that ratings of personality are inherently so unreliable that refinement beyond five or seven scale divisions was probably not justified.

At about this time the quantitative approach to the social psychology of the preschool child was just getting under way. The earlier studies inclined toward rating scales, but the trend soon shifted toward "objective" time-sampling methods. This shift is illustrated by the work of Thomas (1929), Goodenough (1928), and Olson (1929) (23, 13, 18). In recent years, as the limitations of the objective approach have become more and more evident, there has been a swing back toward the rating scale, particularly in some of the more intensive research projects. But by this time Symonds' conclusions had found their way into textbooks and accepted practice (21), with the result that even with elaborate provisions for rater training and careful definition of variables the scales were limited to the relatively crude five- and seven-point form. Here Conrad (1933), Murphy (1937), and Macfarlane (1938) furnish excellent examples (9, 17, 15).

The problem of the optimal refinement of the rating scale has recently been reopened by Champney and Marshall (8), with evidence that under good rating conditions a scale of at least 18 to 24 points may be desirable. For the present it will suffice to point out that Symonds' conclusions were based necessarily upon the low level of obtained reliabilities for rating scales previous to 1925, when progress in developing highly skilled professional raters had scarcely begun. Since that time the seven-point limitation has perhaps rendered the scales too coarse to do justice to further refinement in the skill of the rater.

A second source of undue pessimism regarding rating scales lies in certain misinterpretations of the concepts of <u>reliability</u> and <u>validity</u>. <u>Reliability</u> is the older of the two concepts, and deals with the accuracy of measurement. It has usually been determined by a comparison between repeated measurements. So long as the measurement was a direct one, so long as the yardstick was applied directly to the variable we were trying to evaluate - as is true in the physical measurements such as height and weight - there was no need for subdividing the concept. But when we attempt to evaluate one thing by measuring another, when the measurement is indirect, it is important to distinguish between (a) the accuracy of the measurement itself, and (b) the accuracy with which the measured variable represents the more elusive quality which we are really hoping to get at. This need gave rise to the concept of <u>validity</u>,

which deals with the second point. In actual psychological usage a validity coefficient (unless it is corrected for attenuation) has come to deal with both matters together, whereas a reliability coefficient deals exclusively with the accuracy of the measurement itself.

Thus, when we wish to measure scholastic aptitude, we set up an objective test form employing a variety of informational, reading-ability, and logical-reasoning items which we believe to be symptoms of the aptitude which we cannot readily measure directly. If after many years of trial and revision we secure retest reliability coefficients in the .90's we conclude that the test is sufficiently reliable. But before we are satisfied that it is really measuring scholastic aptitude we correlate it with a measure of actual scholastic achievement. If we can thus obtain validity coefficients between .55 and .65 we are relatively content with our psychometric instrument. Where a measure of achievement is not available we are wont to use other criteria of validity, such as another and better-established objective test, or the combined ratings of experts in judging the quality we are trying to measure.

This is all extremely trite to the psychometrician. The point is that what is comparable to the reliability of a measurement of height or weight, or any direct measurement, is not the reliability of the indirect test, but its validity. And a rating is a direct measurement. For the rater judges the variable itself rather than some objective surrogate. Therefore, if we follow the current practice of taking the correlation between different raters as a measure of reliability, then the reliability of a rating scale should be compared with the validity of a test.

The same end may be reached by a different route. The rating technique may be divided into scale and rater, treating the inter-rater correlation as a validity coefficient, and reserving "reliability of the scale" to refer to correlations between successive ratings made by the same rater (5). In defense of this proposal it has been pointed out that the rating scale is a device for measuring a given rater's judgment, on the assumption that the latter is a fallible index of the true variable as defined. Compared with a test:

Rating Scale : Judgment : True Variable = Test : Objective Sympton : True Variable

It would follow, then, that a correlation with a second rater - or better, with a pool of expert raters - would test the validity of the rating.

Any simple treatment of ratings in terms of the traditional psychometric concepts of reliability and validity leads to confusion unless due attention is given to the peculiarly complex nature of the observer-rater-scale "instrument." This problem will be treated in more detail in the next paper of the series (6). Our immediate concern is to relieve the rating scale of its unearned stigma of hopeless unreliability. Symonds' finding of an average "reliability coefficient" of .55 and an upper limit of .60 or .70 for the earlier attempts at the construction of rating scales (22) takes on quite a new complexion when viewed in terms of validity.

Toward Greater Refinement

In the light of what is known about perception in general, and about the process of judgment in ratings, it does not seem too optimistic to assume that a substantial improvement in the quality of ratings could be achieved by training the rater. With experience comparable to that of other skilled judges - bridge players, music critics, physicians, etc. - professional raters might be developed who could discriminate several times as many degrees of certain complex psychological variables as can the untrained layman. This is particularly important in studies of home conditions, where it is often impracticable to boost the reliability of a coarse measure by combining the ratings of several different observers; and where the single observer, by virtue of general training and experience, and specific acquaintance with the case load and with the scales has an excellent chance to develop a high order of skill.

The problem is to provide a scale to match the skill. The most discriminating judgments will avail us little if they are fogged by a coarse technique of scaling and rating. The variables should be selected with care, and defined so that their meaning is clear to the rater. This clarity of definition should hold not only for the variable as a whole but for various degrees of the variable along the rating line, so that a rating at any given point means the same thing to different raters and to different interpreters of the results. A scaling technique should be employed which provides an appropriate distribution of ratings, avoiding spots in the scale where the ratings will pile up and obscure true differences between homes. The rating procedure and format of the scales should conform to the accepted principles of rating scale practice (12, 3, 21, 14). The scoring scheme should provide for as many discriminations on each variable as the best rater is able to make.[3]

In summary, the problem is to design a rating technique which is adequate to express the integrated judgments of the expert home visitor. Since we may expect discriminations which are considerably finer than those which are ordinarily provided for in rating scales, we have deliberately set out to construct scales which are safely over-refined.

B. THE VARIABLES

The 30 variables of parent behavior were developed in the previous paper (7). They are listed on page 162 below. Although stated in terms of parent-child behavior which may be observed by the home visitor, the variables are far from "objective" in the sense of simple, concrete, specific acts which can be checked off mechanically. Each one is designed to represent a generalized configuration in the child's environment, a configuration which is manifested in a great variety of specific incidents and yet functions as a unitary influence in the child's development. Each variable demands of the rater a complex, and perhaps

[3] *This is subject to practical limits, of course. Conservative statistical standards permit the grouping of data into as few as 24 class intervals without concern for the introduction of appreciable error. Therefore there would be little merit in employing more than 24 working steps in a rating scale. But since the distribution of ratings will differ from scale to scale, from rater to rater, and from sample to sample, 24 working steps can be assured only by providing in the format for a considerably larger number. Then there is the matter of convenience in scoring and computational work. Since single-digit scores are demonstrably inadequate (8), this end is probably well served by limiting the scores to two-digit, positive, whole numbers. The scoring scheme employed is described on page 141.*

largely unconscious, integration of observations and inferences into a quantitative judgment. Each scale must be an instrument by means of which this judgment may be expressed, recorded, and measured.

C. THE GRAPHIC-PARALLEL-VERTICAL FORMAT

A standard form has been adopted to serve for all variables. Each scale is unique only in the verbal material defining the variable, the descriptions arranged along the rating line, and the points at which these descriptions are placed.[4] One complete scale is shown in Figure 1.

The Graphic Form

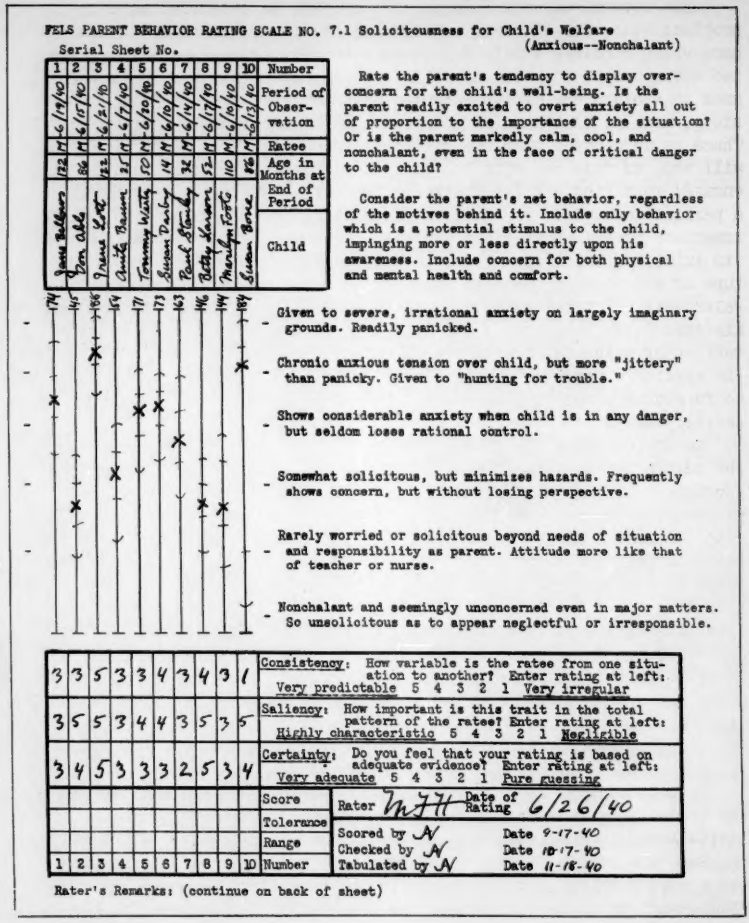
The continuous graphic type of scale seemed to be more appropriate by all odds than the discrete-point type which is so much more widely used. The graphic scale places no limits on precision other than those inherent in the rater and the definition of the variable. The rating line, running without a break from one extreme to the other aids the rater in conceiving of the variable as a smoothly graded continuum. It does require more space on the sheet, and more time to score, than the numerical scale. But these disadvantages can be minimized by efficient design of the blank, and are more than balanced by the greater flexibility of the graphic scale. Scoring stencils for any desired degree of precision may be applied without change in the scale - even to the same completed ratings. Moreover, the spacing of the cues can be finely adjusted along the scale, since they are not rigidly bound to small whole numbers. The graphic scale, furthermore, may derive some motivational advantage from the fact that it is not a constant source of vexation to the conscientious rater when he finds his judgments falling between the defined points.

The Parallel Form

No amount of care in achieving a clear concept of each variable on paper will make the ratings valid unless the niceties of the concept get through to the rater during the rating process. This will be made easier if, once he has established the particular continuum as a frame of reference, he can rate a number of cases before going on to the next variable. Such a procedure is frequently recommended in rating-scale literature as a means of fostering objectivity and minimizing the halo effect (3, 14). But usually the arrangement of the materials, together with a natural inclination (if not necessity) to deal with one case as a unit, prevails over the dictates of the abstract principle even when it is written into the instructions. To help solve this problem we have designed a "parallel-rating" format, in which a separate sheet is devoted to each variable and provides for rating ten cases (see Figure 1).

To take further advantage of the form the rater is encouraged to consider each case not on its intrinsic merits alone but in comparison with the other cases on the sheet. The close parallel arrangement of the rating lines makes such comparisons easy. The descriptive cues are

[4] *The term, "cue," has been appropriated to designate these scaled descriptions, and "cue point," to refer to the short dash preceding the cue and indicating more precisely its position on the scale.*



FELS PARENT BEHAVIOR RATING SCALE NO. 7.1 Solicitousness for Child's Welfare
(Anxious--Nonchalant)

Serial Sheet No.

1	2	3	4	5	6	7	8	9	10	
6/17/40	6/17/40	6/21/40	6/12/40	6/10/40	6/10/40	6/14/40	6/17/40	6/10/40	6/13/40	Period of Observation
M	M	M	M	M	M	M	M	M	M	Ratee
122	86	123	35	150	14	32	52	110	186	Age in Months at End of Period
Jerry Walters	Tom Cole	Donna Lee	Anita Bacon	Tommy Watts	Susan Darby	Paul Steady	Betty Jensen	Marilyn Forte	Susan Bone	Child
174	147	166	169	171	173	163	146	144	189	

- Given to severe, irrational anxiety on largely imaginary grounds. Readily panicked.
- Chronic anxious tension over child, but more "jittery" than panicky. Given to "hunting for trouble."
- Shows considerable anxiety when child is in any danger, but seldom loses rational control.
- Somewhat solicitous, but minimizes hazards. Frequently shows concern, but without losing perspective.
- Rarely worried or solicitous beyond needs of situation and responsibility as parent. Attitude more like that of teacher or nurse.
- Nonchalant and seemingly unconcerned even in major matters. So unsolicitous as to appear neglectful or irresponsible.

Rate the parent's tendency to display over-concern for the child's well-being. Is the parent readily excited to overt anxiety all out of proportion to the importance of the situation? Or is the parent markedly calm, cool, and nonchalant, even in the face of critical danger to the child?

Consider the parent's net behavior, regardless of the motives behind it. Include only behavior which is a potential stimulus to the child, impinging more or less directly upon his awareness. Include concern for both physical and mental health and comfort.

3	3	5	3	3	4	3	4	3	1	Consistency: How variable is the ratee from one situation to another? Enter rating at left: Very predictable 5 4 3 2 1 Very irregular
3	5	5	3	4	4	3	5	3	5	Saliency: How important is this trait in the total pattern of the ratee? Enter rating at left: Highly characteristic 5 4 3 2 1 Negligible
3	4	5	3	3	3	2	5	3	4	Certainty: Do you feel that your rating is based on adequate evidence? Enter rating at left: Very adequate 5 4 3 2 1 Pure guessing
										Score
										Tolerance
										Range
1	2	3	4	5	6	7	8	9	10	Number

Rater MJH Date of Rating 6/26/40

Scored by N Date 9-17-40
Checked by N Date 10-17-40
Tabulated by N Date 11-18-40

Rater's Remarks: (continue on back of sheet)

Fig. 1. Scale for 25th variable, showing standard format, after rating and scoring. (Reduced from 8½ x 11 in.)

thus fortified with something of the flesh-and-blood concreteness of the old man-to-man scales (3), and at the same time the rater can inspect his completed sheet from the rank-order viewpoint as a final check. There is, of course, the possibility that a given case might suffer some distortion due to the particular context of other cases on the sheet, but this danger will be much reduced if not eliminated if we can provide effective cues as fixed points of anchorage.

When several cases are rated at a sitting we encounter another problem: the time elapsing between observation and rating. When a home visit requires a half day it is sometimes as long as ten days or two weeks before the visitor fills out the rating sheets. This would seem to violate one of the accepted principles of observational technique, namely, that what is observed should be recorded immediately. There is no denying that the visitor's impression of a particular home will tend to fade out with the passage of time, especially when he is engaged in making similar visits and appraisals in many other homes - a perfect set-up for retroactive inhibition. But the demand for an immediate record assumes the typical Ebbinghaus curve of retention with its initial steep drop. There is considerable evidence in the literature of the learning process that this type of curve applies only to relatively unrelated details, such as nonsense syllables. English and his associates (11) have shown that with verbal material, whereas retention of verbatim items tends to follow the classic curve, when it is the meaningful pattern or "substance" which is tested there is little or no forgetting, and in some cases improved performance, over a period of several weeks.

An analogy with the present problem is perhaps not too farfetched. The scales are concerned with an integration of many observations into psychologically significant-meaningful-patterns. Too much preoccupation with the concrete details of a fresh home-visit experience will tend to obscure these larger patterns. But after a few hours, or days, the relatively meaningless details - the verbatim material - will have suffered considerable forgetting, leaving in sharp relief the impressions which are important for the ratings. In short, perhaps a time interval is as favorable to a judgment as it is unfavorable to an observation. We have not made a quantitative study of the optimal interval between visit and rating, but the range from one day to two weeks has proved quite satisfactory in the judgment of our Fels home visitors.[5]

The Vertical Form

By some unhappy accident the graphic rating scale got started in the tradition of the yardstick rather than the thermometer. The limitations of format inherent in the horizontal line have effectively dampened any incipient enthusiasm for lengthy cues,[6] or for anything but a most cavalier attitude regarding their spacing along the line. Single-word or short-phrase cues are probably quite adequate for the usual rating technique, where the relatively casual and untrained rater would be impatient and confused by anything else. But brevity is a virtue only when it is a means to effectiveness. Conrad (9) and Macfarlane (15) have found that the trained, professional rater can and will make use of as much as two hundred words of definitional and cue material per variable. If one is interested in exploiting the possibilities of rating-scale precision there is no escaping the need for

[5] *The visitor does write out a brief report immediately after each visit. Although it does not cover the same ground as the ratings, dealing largely with anecdotal items, changes in home personnel, living arrangements, etc., rereading these notes probably helps to bring back the impressions growing out of the visit.*

[6] *See earlier footnote, p. 137.*

more verbal material than can be squeezed into the horizontal format. Perhaps this was one reason why Conrad preferred a point scale. Macfarlane not only uses a point scale, but prints the verbal definitions and cues in a separate manual, a practice which saves space but probably places some additional burden on the rater.

By standing the rating line on end the problem is solved rather nicely. More cue material can be used, and it can be arranged with maximal legibility directly on the rating sheet. The unbroken continuum of the graphic scale can be retained, and yet the scale position of the cues can be readily adjusted to meet the requirements of a sophisticated scaling technique. (See Figure 1)

In the parent scales we have employed from fifty to a hundred words for a general definition of the variable, followed by from five to seven cues along the rating line. The cues contain from four to twenty-five words each. The exact scale position of each cue is made clear by the cue point, placed between the cue and the rating line. A second set of cue points is placed at the extreme left of the sheet, to aid the eye in projecting the cue point accurately across the ten lines. The cue points are intentionally kept out of the actual rating-line area so as not to break the continuum into discrete steps.

Clearly a graphic scale is only pseudo-graphic if the rater is constrained to place his mark at one of several predetermined points on the line. The mere presence of cues tends to act as a partial constraint of this sort, especially with careless or untrained raters. The resulting "saw-toothed" distribution of ratings, with minor modes over the cues, may be observed in the typical graphic-rating histogram (12). Such distortion may be a serious matter where the cue modes are sharp and the cues are far apart - the graphic scale then tends to become a point scale with as many scale points as there are cues. In any case, the tendency to rate at the cues operates to reduce the accuracy of the rating as an expression of the rater's best judgment. We have attempted to avoid this source of error in two ways, (a) by keeping the rating lines clear, and (b) by specifically warning the rater against the tendency (see instructions on page 163). Emphasis is placed on cultivating a concept of the scale as a smoothly graded continuum, defined by all the verbal material taken together. One function of the single cue is to add definitive color to the entire scale. Its specific function as a meaningful anchor at the cue point, vital as it is to the validity of the scale, should not be permitted to lead the rater to treat each cue as if it were a multiple-choice item, or even a complete definition of the cue point.

In a vertical scale it is by all means more suitable to use the upper end of the line to represent the high degree, large amount, or positive aspect of the variable as named and defined. This arrangement has been followed throughout, despite the fact that many authorities have advocated an alternation or a randomization of the polarity of the several scales in a battery as a means of breaking up the halo effect (3, 21, 14). Our scales are descriptive rather than normative; nothing is stated or implied as to what constitutes <u>desirable</u> behavior. Although in many of the scales the point which most people would consider "ideal" is probably nearer the top than the bottom, there are several, such as

Discord and Disciplinary friction, where the opposite is the case. Rather than introduce confusing reversals in polarity we have thought it best to rely on the training of the rater, the parallel-rating format and procedure, and the objectivity of the definitions in combatting halo tendencies.

Scoring

The length of the rating line has been set at 90 millimeters, or about three and a half inches. This fits the requirements of the format, and is not too different from the usual "four to five inches" (21). Its chief virtue, however, lies in its adaptability to efficient scoring and subsequent statistical treatment. The score is defined by dividing the line into ninety units, calling the bottom unit "10" and the top, "99." This is in keeping with the "over-refinement" objective discussed earlier (page 136), making maximal use of the two-digit range and avoiding the confusion which would be invited by extending the range below 10. If, for some purposes, it is desirable to use coarser scores, by simply ignoring the second digit we have a nine-point scale running from "1" to "9" in centimeter units. Either the millimeter or the centimeter scores are admirably adapted to statistical handling by modern calculating machines or by punched-card methods. In practice the scoring can be done very simply by applying an ordinary millimeter stick to the rating line so that the bottom end of the line reads 9.5 mm., and reading the rater's mark to the nearest millimeter.

D. SUPPLEMENTARY RATINGS

Despite a high degree of both motivation and training, various raters will differ in the meticulousness with which they approach the rating task. In an effort to encourage a deliberate and painstaking attitude, to standardize the rater's procedure, and to permit expression of nuances of judgment not taken care of by the primary rating, we have provided for five supplementary ratings. Three of these, Consistency, Saliency, and Certainty, are incorporated in the format (see Figure 1). The other two, Tolerance and Range, call for additional marks on the rating line. All five are described in detail in the instructions to the rater (see page 163).

The data from the supplementary ratings may be found useful in various ways. Conrad, in his California Behavior Inventory (9), employs two supplementary judgments, which he finds to be useful in selecting the more reliable ratings (10). Then, certain of the supplementary scores may give us additional information about the home, such as the consistency of parent behavior, or the relative importance (saliency) of various characteristics in the total profile. The graphic ratings of tolerance and range may be directly useful in case studies without the need for any scoring or computations. Inspection of the original rating sheet thus not only tells us, in terms of the descriptive cues, what level of behavior occurs characteristically in this parent, but over what extreme range his behavior varies from one situation to another.

But useful as the supplementary ratings may prove in individual instances, they are not an essential part of the technique. They may be thought of as optional adjuncts - part of the campaign to squeeze the last drop of discrimination from the human instrument. Where the rater's habits of systematic devotion to the task have been well established, or where there is little concern for precision or for the supplementary data, considerable burden on the rater may be lifted by their omission.

E. SCALING PROCEDURE

The heart of any rating-scale technique lies in the verbal material which defines the scale. No trick format, fancy scoring scheme, or painstaking procedure will make up for a failure to convey to the rater a sharp concept of the variable he is rating. He must know not only the general nature of the variable, but also the status of closely related concepts - are they to be included in the evaluation, or disregarded?

But a clear picture of the variable as a whole is not enough. It is not enough that he understand the quality which is high at the top end of the scale and low at the bottom; he must have a clear picture of the kind of concrete behavior which exemplifies the variable at various points along the scale. He must have all these things and have them consistently from time to time and from case to case. It is the function of the verbal material to convey these things consistently, not only to him but with constant meaning to all persons using or interpreting the scales.

This is a large order; in fact, when stated in this perfectionistic form, it is too large for the limited capacity of human language. The miracle is that the job can be done at all! In the quest for precision in rating-scale technique this problem would seem to be the ultimate bottle neck: the effectiveness of the scale in conveying a precise and constant meaning sets the limit of refinement. Here then is the place to spare no pains.

The scaling procedure falls into three parts which can be treated separately, namely, the writing of the general definition of the variable, the writing of the cues, and the determination of a suitable scale position for each cue. These will be taken up in order.

The General Definition

Much of the preliminary definitional material was already at hand as a byproduct of the procedure followed in selecting the variables - described elsewhere (7). This material was rewritten with a view to concise diction and concrete description of the behaving parent. An effort was made to anticipate, or to discover by trial-and-error, the various perplexities of interpretation which would occur to the rater in actual service, and to clear them up in advance. To what kinds of situations does the variable apply? Is an allowance to be made for age of the child? Are we to rate what the parent feels, or what he does? Do we consider the behavior of the parent regardless of the whereabouts

of the child, or only what happens to the child? Et cetera. It was found helpful to employ a fairly standard form for the general definitions, roughly as follows:

> Rate the parent's tendency to do so-and-so. Does he typically do so-and-so (describing behavior in the upper part of the scale)? Or does he do so-and so (describing the lower end)?
>
> This variable applies only to so-and-so situations. Disregard so-and-so. Allow for so-and-so. Include so-and-so and so-and-so. Do not confuse with so-and-so.

These paragraphs are preceded by a name for the variable and a brief indication of the high and the low ends of the scale, as "Readiness of Enforcement (Vigilant-Lax)." Although the name was adopted reluctantly, as a necessary means of identification and reference only, there is no doubt that it plays an important part in the functioning definition, especially to persons interpreting results and lacking a background of work with the complete scales. This is unfortunate, because in spite of our best efforts we were unable to find names which in all cases adequately represented the variable. It cannot be too strongly emphasized that the mere name of a scale may be very misleading: the variable can be understood only by reference to the entire scale.

Criteria for a Good Cue

Cue writing, like the selection of variables, discussed in a previous paper (7), is a creative process akin to the formation of hypotheses. The final check lies in the empirical test of performance: If the cues do not do their job well - notwithstanding adequate definition - the scales will lack validity, and the ultimate picture of parent-child relationships will be blurred. Out of our trial-and-error experience with parent-behavior cues certain guiding principles have emerged which can be summarized in terms of a list of ideal specifications for the writing of cues. Since some of the requirements tend to be mutually contradictory it is usually impossible to satisfy all of them, and the end product must be viewed as an attempt at an optimal compromise.

1. Clarity. The cue should be short, simple of vocabulary and sentence structure, and unequivocal in its meaning. It should be easily read and comprehended by the rater.

2. Relevance. The described behavior should fit into the variable defined, and should be as free as possible from determination by other variables.

3. Broad scope. Some variables are composed of closely related sub-variables. General Babying, for example, breaks down into physical and social forms; and most variables permit of sub-categories of frequency and intensity. Each cue should ideally relate to all the sub-variables. This is usually impossible, but the same end can be approximated by treating different aspects of the broader variable in different cues.

4. Narrow scale range. A good cue applies only to a relatively short distance along the scale, avoiding overlap with adjacent cues above and below. This is especially difficult to achieve in the middle range. Sometimes a double-barrelled statement will turn the trick, as in "Helps when needed, but not when child can get by alone."

5. Incisiveness. A cue must be more than mere words. It should describe behavior with as much concrete vividness as is compatible with the breadth of the definition. The use of words like "rarely," "usually," "slightly," and "extremely" is only excusable if the scale value of the cue does not depend on them. The incisive cue will convey to different persons the same meaning in terms of actual behavior.

6. Variety. Repetition of the same words and phrases from cue to cue tends to confuse the reader. A variety of expressions and types of situations, if they are well chosen and relevant, helps achieve breadth and incisiveness in the scale as a whole.

7. Objectivity. The scales should be kept free of any notion of goodness, rightness, or wiseness, or any other connotation of social norms. What varies along the scale from high to low must be an intrinsic quality of parent behavior conceived in terms of psychological stimulus to the child. Laudatory and derogatory terms must be carefully avoided. If there is such a thing as an ideal or a correct score we do not know where on the scale it falls.

The same taboos apply to cues suggesting distributional norms, such as "average," "below average," and "exceptional." These are almost as destructive to research as are "poor," "good," and "excellent." We shall not learn much about what constitutes either the ideal parent or the average parent unless *a priori* assumptions about these things are ruled out of our yardstick.

8. Number of cues. In constructing the present scales we have started with anywhere from 5 to 21 preliminary cues, and have retained at least 5 and not more than 7 in the final form. Extremes are described with relative ease. To construct incisive cues for the in-between points is more difficult, but essential if the scale is to have any internal stability. Three cues, one near each end and one near the middle, should be an absolute minimum for any graphic scale. In certain variables where there is a readily defined center between two opposites, as in Approval-Disapproval, a minimum of five cues is called for.

9. Spread. Unless the scale provides an adequate dispersion of cases it fails in differentiating one case from another as well as it might. We have attempted to avoid dead ends by tempering the extremes sufficiently to keep them within the limits of reality. No cue should be so extreme that no conceivable case could answer that description. In regions of the scale where a piling up of cases may be expected (often the center), it is well to insert more cues and keep them well spaced.

10. Independence of child's age. The specific behavior through which a given level of a parent variable is expressed will sometimes change greatly with the increasing age of the child. "Ties child's shoes" represents a higher degree of babying for a six-year-old than for a child of two. The use of such cues would require age norms. In attempting to avoid this we have aimed at definitions and cues which are relatively independent of child's age. Babying, for instance, is treated as "helping the child more than necessary," the assumption being that where the basic parental tendency remains constant the child will get as much uncalled-for help, and therefore the same rating on the scale, at twelve as he did at three.

Cue Scaling by the Graphic Method

The two processes of cue writing and cue scaling are not entirely separable. The scaling procedure serves as a partial check on the adequacy of the cues. If one or more cues in the first set turn out to be ambiguous in scale value they will fail to survive the scaling procedure. In this case the cues are revised or rewritten entirely and scaled again. This process is repeated until a set of cues is obtained which meets the criterion of satisfactory scaling. In case several revisions fail to show progress toward an acceptable scale the variable itself probably needs redefining. The graphic method of cue scaling proceeds as follows:

1. The preliminary cues are typed, each on a three-by-five card, and shuffled. Each card is labeled with a distinctive but non-serial symbol, such as the signs found on the typewriter (#, $, %, &, *, ©, ¢, ?, etc.). A definition card is prepared, giving the name of the variable and the general definition.

2. Five or more judges are selected, preferably having about the same degree of familiarity with the subject matter of the ratings as the persons who are to use the scales. A rating card is provided for each judge, with a 90-millimeter line along which he may arrange the cue symbols.[7]

3. The cards are presented to each judge, independently, with the following instructions:

> Please read over carefully the definition card and the cue cards. Now sort the cue cards into rank order, giving top rank to the description belonging highest on the scale defined. Then, to express more accurately your judgment of scale values, place the symbol for each card in the appropriate position along the graphic line on the rating card. Any remarks or suggestions about items may be written on the rating card.

4. The rating cards of the several judges are then lined up in parallel and inspected. Unless at least five judges are unanimous in the rank order assigned each cue, and unless the cues have been accorded a reasonable spread along the line, the cues are revised in the light of the results and the entire procedure repeated. When the results do meet this criterion of satisfactory scaling the final scale is drawn up,

[7] *The writer is indebted to the following members of the Fels staff for their patient and invaluable services in the role of judge: the Misses Katherine Kampp, Helen Marshall, Jeannette McGraw, Helen Newbery, and Harriet Smith, and Drs. Virginia L. Nelson, T. W. Richards, and L. W. Sontag.*

assigning each cue to a position approximating the average position given to it by the judges.[8]

Figure 2 illustrates the graphic cue-scaling method through two revisions to the final cues for Solicitousness, the same scale shown in Figure 1.

The matter of revising the cues after each preliminary attempt at scaling deserves some further comment and caution. Cues which are given widely different positions by various judges should, of course, be eliminated as ambiguous. Where reversals occur in adjacent cues they can sometimes be combined to advantage, or one of them omitted. Sometimes the changing of a word or two will suffice to alter the value of a cue in a desired direction. Comments about individual cues collected from the judges are often very helpful. But above all it is essential that at all stages of revision the criteria for a good cue be kept in mind. The scaling process is a check on some of these criteria, notably, narrow scale range. But it is no check at all on the most important criteria, incisiveness and objectivity.

There is a real danger that the cue-scaling procedure places too much emphasis on the relative scale value of the cues and not enough on their effectiveness as anchor points. Certain types of cues, all too common in rating-scale practice, readily yield unanimous judgments as to rank order, namely, abstractions such as "scarcely--moderately--extremely." Such cues serve little purpose. They point to no agreed-upon levels of behavior. They may help an unskilled rater to put a mark on a line, but they do not convey to anyone else what that mark means. And they are no help at all to the expert, who already knows that the rating line itself stands for a range from "scarcely" to "extremely."

What is needed is a verbal characterization of behavior which will indicate what is "moderate," "extreme," etc. A good cue must cut through the superficial level of verbal abstraction and anchor itself solidly to a distinctive spot in the underlying continuum of flesh-and-blood environment. Incisiveness must not be sacrificed for the sake of easy scaling and its check on mere relative scale value. It is quite possible that a vivid and effective set of cues might baffle the judges when presented in shuffled form and yet "click" beautifully when the cues appear in their assigned positions along the rating line. But the unanimously sorted abstractions are all too likely to prove to be a series of will-o'-the-wisps. Since the scaling process provides no check for incisiveness, it is important throughout the rewriting and revising to reject all cues which are mere empty verbal distinctions, regardless of agreement among judges as to their scale position.

[8] *In the scaling of cues for the Fels Parent-Behavior Scales it never required more than three revisions to meet the criterion, and usually two were sufficient.*

The graphic method of cue scaling as described here is the end product which evolved as a technique as more and more variables were scaled. Only the last six variables to be scaled

First Edition of Cues

& Shows worry over the most trivial matters. Often suffers acute, irrational anxiety on imaginary and insufficient grounds. Readily panicked.

* Exaggerated vigilance; always "hunting trouble." Chronic anxious tension over child, but more inclined to "jitters" than to panic.

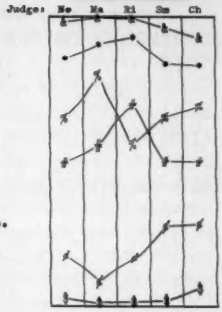
% Definite tendency to be over-anxious when child is in any danger, but usually keeps concern under rational control.

Alert to dangers. Shows anxiety with due cause, but without losing perspective

/ Rarely worried or solicitous beyond minimal responsibility as parent and the needs of the situation. Attitude more like that of teacher or nurse.

(Nonchalant and seemingly unconcerned even in major matters. So unsolicitous as to appear neglectful or irresponsible.

Graphic Judgments



Second Edition of Cues

* Worried over trivialities, irrational anxiety on largely imaginary grounds. Readily panicked.

& Over-solicitous; "hunts trouble"; chronic anxious tension over child, but more "jittery" than panicky.

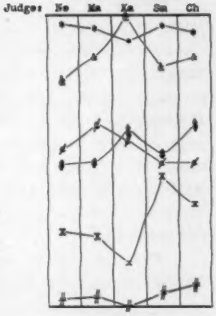
(Tends to show some anxiety when child is in any danger, but usually keeps concern under rational control.

/ Reasonably solicitous, but minimizes hazards. Shows anxiety with due cause but without losing perspective.

x Seldom worried or solicitous beyond needs of situation and responsibility as parent. Attitude more like that of teacher or nurse.

Nonchalant and seemingly unconcerned even in major matters. So unsolicitous as to appear neglectful or irresponsible.

Graphic Judgments



Final Edition of Cues

/ Given to severe, irrational anxiety on largely imaginary grounds. Readily panicked.

% Chronic anxious tension over child, but more "jittery" than panicky. Given to "hunting for trouble."

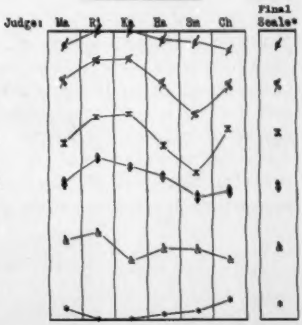
x Shows considerable anxiety when child is in any danger, but seldom loses rational control.

(Somewhat solicitous, but minimizes hazards. Frequently shows concern, but without losing perspective.

& Rarely worried or solicitous beyond needs of situation and responsibility as parent. Attitude more like that of teacher or nurse.

* Nonchalant and seemingly unconcerned even in major matters. So unsolicitous as to appear neglectful or irresponsible.

Graphic Judgments



* The final scale will be found in Figure 1.

Fig. 2. Cue scaling by the graphic method — Scale No. 7.1 solicitousness.

THE FELS PARENT-BEHAVIOR RATING SCALES

To save space, the final battery of 30 scales is being presented in text form, the unique portion of each scale being given but the spacing and cue points omitted. It should be noted that the spacing of the cues along the scale is unique for each scale. The standard format for the scales may be seen in Figure 1. For the actual spacing of cue points the reader is referred to the scales themselves (4).

Scale No. 1.1 Adjustment of Home (Well-adjusted—Maladjusted)

Rate the general internal adjustment of the family as a whole in its day-by-day relationships. Is the home atmosphere characterized by satisfaction, stability, achievement, and happy adjustment; or by thwarting, unpleasantness, repression, and insecurity?

This is a broad variable, including conflicts among persons, among motives, or with obstacles. Conflicts may be emotional, social, economic, or physical; and both overt and covert. Rate the total configuration—the quality of the child's home atmosphere.

Exceedingly well-adjusted. Characterized by pleasant co-operation, security, and full satisfaction throughout.

Fundamentally sound adjustment, but with minor conflicts here and there.

Fairly smooth on surface, but suggests undercurrent of repression or insecurity.

Definite evidence of mild maladjustment throughout.

Dominated by maladjustment, coloring most family activities.

Extreme maladjustment; torn with conflict, repression, and insecurity.

Scale No. 1.2 Activeness of the Home (Active—Inactive)

Rate the general activity level of the home, taking the household as a whole. Is the home atmosphere active, quick, and alert; or is it inactive, slow, and inert?

This is a broad, general variable, including amount and quickness of activity, alertness, decisiveness, and tension, insofar as they are manifest overtly as part of the child's environment.

employed the full graphic method (Nos. 3, 5, 18, 23, 25, and 26); the others were scaled on the basis of rank order judgments, employing the same number of judges and the same requirement of unanimity, but the final spacing was in terms of the writer's judgment alone. The graphic method was found to be little additional trouble to the judge, and much more helpful in revising and final scaling, actually simplifying the work more than it complicated it. The graphic method, substantially as described, was used with good results in a subsequent scaling project dealing with nursery-school ratings of child personality (20).

It may be noted that our method of cue scaling bears more than a superficial resemblance to the method of equal-appearing-intervals devised by Thurstone for the scaling of social attitudes (24). We do not, however, make any claim of equal scale units in the psychophysical sense. As Boring has shown (2,1), neither equal scale units nor a normal distribution are worthy goals of a good psychological scale. The present approach has aimed, rather, at a scale tuned to reflect the best rater's ability to discriminate between cases - a goal which is probably sufficiently close to equal units and normal curves for all practical purposes. These issues are discussed in some detail in the original thesis, pages 55-62 (5).

Home extremely bustling, busy, excited, tense.

People in home move quickly, talk rapidly, work with dispatch. Home alert, wide-awake, moving, decisive.

People move, talk, and work without haste, but with some dispatch. Home alert, but not hypertense.

People move, talk, and walk with leisurely deliberateness. Home relaxed, but not lackadaisical.

People move slowly, talk slowly, work slowly. Home passive, relaxed, easy-going, indecisive.

Home poky, lackadaisical, lazy, slow-moving, procrastinating.

Scale No. 1.5 Discord in the Home (Conflict—Harmony)

Rate the extent to which the home surrounds the child with an atmosphere of overt conflict, discord, unpleasant argument, recrimination, quarreling, complaining. Is the home atmosphere marked by unpleasant discord among individuals? Or are the inter-personal relations of the household typically harmonious, friendly, and agreeable?

Disregard conflict between the object-child and other individuals. Include discord among siblings, domestics, parents, etc. Disregard other types of maladjustment, as worry, insecurity, grief, illness, cynicism. Include discord only insofar as it impinges more or less directly upon the child.

Household flies into vindictive recriminations, bitter disputes, on slightest provocation.

Underlying discords often break through the general surface harmony as sharp arguments, or ill-natured sarcasm.

Frequent unpleasant wrangling, complaining, squabbling.

Harmonious basic relationships, overlaid with a good deal of surface contention, bickering, and teasing.

Tolerant, friendly. Arguments tend to be good-natured. Teasing occasional. Quarrels rare.

Peaceful, harmonious, agreeable atmosphere reigns. Household harmony disturbed only under rare and extreme circumstances.

Scale No. 1.6a Sociability of Family (Aggressive—Reclusive)

Rate the family as a whole according to the degree to which its energies are directed outward from the home toward society. Is the family constantly initiating contacts and participation with neighbors, friends, relatives, and in the church, lodge, P.T.A., etc.; or is it reclusive and inwardly oriented, taking little interest in and avoiding contacts with the community?

Include all social contacts outside the immediate household, excepting necessary business routine. Include both breadth and depth of sociability. Include interest, activity, initiative, hospitality.

Family constantly active socially; always seeking new contacts; eager and uninhibited mixers.

Actively interested in what is going on outside the household; mix readily, without being over-aggressive.

Accept contacts with enthusiasm, but hesitate to intrude into the lives of others.

Neither avoid nor initiate social activities; accept contacts passively.

Tend to avoid direct social intercourse; seldom go out; rarely entertain.

Family resents social advances from outside the household; never mixes socially.

Scale No. 1.7 Co-ordination of Household (Co-ordinated--Chaotic)

Rate the routine functioning of the household as to its smoothness of organization. Is it effectively planned and executed? Or is it uncoordinated and chaotic?

Rate on basis of effectiveness in operation rather than tendency to systematize every detail. Include care of belongings, co-ordination of schedule, planning, meeting of responsibilities, and general efficiency of organization as it works in practise. Disregard variations in aesthetic standards, style, form, socio-economic status, etc.

Extremely effective management. Model of efficiency. Long-range planning, flexibly executed. Confusion unknown.

Smooth-running and efficient on the whole. House kept in order and on schedule most of time. Meals, finances, education planned ahead. Some superficial disorder.

Fair co-ordination. Considerable disorder, but can usually find things. Buying inefficient, but meals fairly adequately planned. Sometimes off schedule, but never miss trains.

Poor co-ordination. Essential rudiments of organization are there, but inefficiency and confusion are common. Often late; off schedule half the time. House disorderly.

Chaotic, disorganized. Nothing happens on schedule. No planning. Equipment in tangled scramble. Confusion reigns even in essentials.

Scale No. 1.91 Child-centeredness of Home (Child-centered--Child-subordinating)

Rate the organization of the household according to the degree to which it is built around the child's needs and welfare. Does the child get more than his proportionate share of consideration; or is his welfare subordinated to that of other members of the household?

Rate for the specific child, as against the entire remaining household including siblings. Behavior is "child-centered" to the extent that it involves sacrifice of pleasure, convenience, opportunity, etc., in attempting to benefit the child.

Whole household revolves around child; many major sacrifices for child's trivial comforts.

Consideration for child clearly predominates, but not to the exclusion of other interests.

Child's welfare gets slightly more attention than the welfare of others.

Child gets proportional consideration; is as often disregarded as sacrificed for.

Although given attention in critical matters, on the whole, child gets neglected in favor of other interests.

Household organized around interests of other members. Child

definitely neglected even in essential matters.

Scale No. 2.11 Duration of Contact (Extensive--Brief)

Rate the parent in terms of the number of hours (for an average day) of actual contact with the child.

"Contact" is here defined as a situation in which parent and child are sufficiently alert and close together to make social intercourse possible--amount of opportunity for parent-child stimulation.

Base rating on total impression, from observation and conversation. Avoid direct quizzing as basis for rating. Judge in terms of average for a typical week. (To correct weekday figure, add 1/7 of week-end excess contact.)

Entire waking day together.
Eight hours per day together.
Four hours per day together.
Two hours per day together.
One hour per day together.
Half hour per day together.
Quarter hour per day together.

Scale No. 2.12 Intensity of Contact (Vigorous--Inert)

Rate the reactivity of the parent during contacts with the child. Does the parent react readily and vigorously; or does the parent tend to disregard the child during contact situations?

"Reactivity" includes both initiating social intercourse with the child and responding to the child's initiative. It includes attention, suggestion, affection, coercion, help, conversation, criticism, information, play, scolding, threatening, feeding, explaining, etc.

Rate only on situations where there is opportunity for stimulation--independent of duration of contact.

Intensely vigorous, over-stimulating, excited.
Active, readily attentive, vigorous.
Fairly active, responsive, alert.
Accessible, interested, half-hearted, reserved.
Perfunctory, passive, retiring, taciturn, bored, busy.
Oblivious, absorbed, inaccessible, preoccupied.

Scale No. 3.11 Restrictiveness of Regulations (Restrictiveness--Freedom)

Rate the restrictiveness of the regulations set up or implied by the parent as standards to which the child is expected to conform. Are the requirements numerous and severe; or few and mild? In meeting these standards would the child be highly circumscribed in his behavior, or would he still have a large measure of freedom?

Disregard whether requirements are sharply codified rules, or merely implied in the disciplinary policy. Disregard the parent's motives, and methods of enforcement. Include both prohibitions and positive requirements. Consider the standards <u>expected</u> regardless of how well they are enforced. Rate relative to child's age.

Parent's standards for child's conduct are minutely restrictive beyond all reasonable interpetation of either child's welfare or

family convenience.

Requirements are unnecessarily abundant and exacting, but usually aimed at practical ends rather than "pure discipline".

Restrictions are moderate and practical, but parent shows little concern for child's freedom as an end, slapping on requirements whenever they seem expedient.

Standards and regulations are somewhat liberal. Freedom is allowed in a few matters commonly subject to regimentation.

Child is expected to conform to a few basic standards, but parent will endure considerable annoyance rather than unduly restrict child's freedom.

Standards are both scarce and mild, limiting child's freedom barely enough to avoid the police and the hospital.

Scale No. 3.12 Readiness of Enforcement (Vigilant--Lax)

Rate the parent's tendency to enforce the standards of conduct set up for the child. Does the parent follow up to see that the child conforms, or sustains a penalty? Or are lapses in compliance disregarded?

This variable applies only to situations where there is an opportunity for the parent to enforce an accepted standard which has been, or is being, or is about to be violated by the child. Disregard the methods of enforcement and the severity of penalties. Disregard effectiveness of enforcement, and clarity to the child of standards involved. Do not confuse with the non-regulational type of parental domination covered by the "Suggestion" scales.

Eternally vigilant. Goes out of way to discover and discipline misconduct. Often pounces before lapse occurs.

Seldom lets child "get away with anything." Enforces rules strictly whenever violations come to attention, but seldom deliberately hunts for misbehavior.

Moderately firm. Strict about important requirements and prohibitions; but rather lax with minor violations, especially when they are not an issue at the moment.

Reluctant to enforce standards. Tends to overlook violations unless they are flagrant, cumulative, or threaten serious consequences.

Extremely lax. Disregards obvious misbehavior. Enforces regulations only when pressed by the strongest motives or the severest circumstances.

Scale No. 3.13 Severity of Actual Penalties (Potent--Weak)

Rate the severity of penalties imposed when parent takes official note of misconduct. Are penalties acutely severe, or light and inconsequential?

Do not consider situations where parent entirely disregards misconduct, invoking no penalties. Include all censorial reactions from mild verbal reproof to severe corporal punishment and removal of privileges. Consider only such situations as come under express or implied regulations and standards. Try to judge the penalties in terms of their negative motivating power for the particular child.

Severe penalties, frequently stimulating child to dread, terror, or deep personal resentment.

Rather severe on the whole, but inclined to be lenient in extenuating circumstances.
Moderate penalties. Severe enough to have definite motivating power for the child; but not so severe that the child is over-inhibited, severely frightened, or deeply resentful.
Mild penalties predominate. May be severe in critical situations; but penalties often seem too mild to have much motivating power.
Most flagrant misbehavior provokes no penalty more severe than weak verbal remonstrance. Penalties are so light that their potency for the child is negligible.

Scale No. 3.14 Justification of Disciplinary Policy as Presented to the Child (Rational—Arbitrary)

Rate the parent's tendency to explain to the child the reasons for requirements and penalties. Does the parent attempt to put all discipline on a logical basis? Or are his policies presented in purely arbitrary fashion to the child?

Disregard restrictiveness of regulations, and readiness and severity of enforcement. Disregard the clarity with which regulations are codified, and the extent to which they are democratically set up. Include all control measures, whether pertaining to established policies or merely involving immediate suggestion.

Goes out of way to show child practical reasons behind requirements and suggestions, even in emergencies or where explaining is difficult.
Attempts to explain policies to child, as a general rule, but frequently arbitrary where the issue is very critical or complex.
No apparent tendency favoring either the peremptory or the rational approach to child control.
Arbitrary in most matters. Does not justify policies unless pressed. Often avoids the issue, or invokes moral precepts as reasons.
Never explains policies to child. Handles discipline in very arbitrary fashion, expecting child never to question "why."

Scale No. 3.15 Democracy of Regulation and Enforcement Policy (Democratic—Dictatorial)

Rate the parent's tendency to share with the child the formulation of regulations for the child's conduct. Does the parent give the child a voice in determining what the policy shall be? Or does the parent hand down the established policy from above?

Disregard immediate issues not covered by policy (see Coerciveness of Suggestion). Rate independent of justification of policy to child, and independent of restrictiveness of regulations. Include both overt consulting with child and considering child's expressed wishes. Dictatorial policies may be wise or foolish, benevolent or selfish.

Endures much inconvenience and some risk to child's welfare in giving child large share in policy forming. Consults with child in formulating policies whenever possible.
Attempts to adjust policies to child's wishes wherever practicable. Often consults child.
Deliberately democratic in certain safe or trivial matters, but

dictates when there is a sharp conflict between child's wishes and other essential requirements.

Neither democratic nor dictatorial, deliberately. Follows most practical or easiest course in most cases.

Tends to be rather dictatorial, but usually gives benevolent consideration to child's desires. Seldom consults child.

Dictatorial in most matters, but accedes to child's wishes occasionally when they do not conflict with own convenience or standards.

Dictates policies without regard to child's wishes. Never consults child when setting up regulations.

Scale No. 3.16 Clarity of Policy of Regulations and Enforcement (Clear—Vague)

Rate the clearness with which the parent's standards of child conduct are manifested to the child. Are regulations and requirements clearly formulated and consistently executed, so that the child should be able to know what is expected of him and what will happen if he fails to conform? Or are the parent's standards and policies so vague or fluctuating that the child has little chance of adjusting?

Schedule and other standards are precisely formulated and adhered to meticulously. Parent goes out of way to maintain clear, consistent policy regardless of special circumstances.

Policies are sometimes adjusted to meet unusual circumstances, but on the whole they are clear-cut and consistent.

There is a core of reasonable consistency about parent's policy, which serves as a stable basis for adjustment despite numerous minor fluctuations and vagueness about details.

Standards are usually formulated, but exceptions and modifications are frequent enough to keep child readjusting. Schedule often upset.

Regulations vaguely formulated. Enforcement uncertain and inconsistent. Child's basis for adjustment is slight, even in some major matters.

Policies of parent in dealing with child are so erratic, unformulated, and inconsistent that child can never know what to expect. Schedule chaotic.

Scale No. 3.17 Effectiveness of Policy of Regulations and Enforcement (Successful--Unsuccessful)

Rate the degree to which the child's conduct meets the standards set by the parent. In the eyes of the parent, is the child well-behaved? Or does the child fail to meet the requirements implied in the parent's control policy?

Rate in terms of the child's net overt behavior, disregarding the amount of coercion, threats, penalties, etc., employed in producing the behavior. Disregard child's inner conflicts in conforming.

Child conducts himself in accord with the parent's standards in every respect. Parent's policy achieves its goal.

Parent attains goal in all major respects and most minor ones.

Policy predominantly successful, although it fails in many instances

and respects.

Successfulness very questionable. Parent's regulations and enforcement fail to produce the desired results about as often as they succeed.

Parent's policy fails to elicit the desired behavior in most of the important aspects of control.

Child's overt behavior is entirely at odds with standards implied in policies of parent. Policy completely unsuccessful.

Scale No. 3.18 Disciplinary Friction (Contentious—Concordant)

Rate the amount of overt parent-child conflict over the enforcement of regulations and requests. Is the relationship between parent and child characterized by continuous wrangling, resistance, and rebellion in regard to child conduct? Or is discipline characterized by harmonious co-ordination, without bickering, threats, refusals, and penalties?

Disregard whether child's conduct meets parental standard. The question is, how much disharmony occurs as part of the parent's attempt to control the child's conduct, both in enforcing routine standards and in making immediate suggestions.

Situations to which regulations or standards apply are always characterized by overt parent-child conflict. Parental demands resisted. Friction continuous and acute.

When child is supposed to do (or not to do) something, there is usually an argument, struggle, threat, or penalty. Friction frequent and marked.

Parent invokes penalties, child resists, etc., rather frequently, but harmonious adjustment in disciplinary situations is somewhat more usual. Friction moderate.

Parent-child clashes occur now and then, but they are exceptional, superficial, or mild.

Disciplinary conflicts are exceedingly rare. Either the child conforms docilely, or the parent tranquilly permits lapses. Friction extremely mild or absent.

Scale No. 3.21 Quantity of Suggestion (Suggesting—Non-suggesting)

Rate the parent's tendency to make suggestions to the child. Is the parent constantly offering requests, commands, hints, or other attempts to direct the child's immediate behavior? Or does the parent withhold suggestions, giving the child's initiative full sway?

This does not apply to routine regulations and their enforcement. Rate only where there is opportunity for suggestion. Note that "suggestion" is defined broadly, including direct and indirect, positive and negative, verbal and non-verbal, mandatory and optional.

Parent continually attempting to direct the minute details of the child's routine functioning, and "free" play as well.

Occasionally withholds suggestions, but more often indicates what to do next or how to do it.

Parent's tendency to allow child's initiative full scope is about equal to tendency to interfere by making suggestions.

Makes general suggestions now and then, but allows child large measure of freedom to do things own way.

Parent not only consistently avoids volunteering suggestions, but tends to withhold them when they are requested, or when they are the obvious reaction to the immediate situation.

Scale No. 3.22 Coerciveness of Suggestion (Mandatory--Optional)

Rate as to their dictatorial quality the parent's suggestions in dealing with the child's immediate behavior. Does the parent attempt to control a situation by issuing orders or commands, to be obeyed? Or does the parent make his suggestions optional, or discretionary, with the child?

Apply only where parent is trying to influence child. Try to see through the verbal form to the significant content for the child—does the suggestion demand obedience, or is it a "mere suggestion?"

Efforts to control child take form of peremptory orders, to be obeyed at once, even in trivial matters.

Suggestions not quite absolute in coerciveness, but immediate compliance is expected in matters of any importance.

Parent coercive in major affairs, but uses optional suggestions where there is no important issue.

Definite tendency to avoid coercion where possible, but uses it when exasperated or persistently unsuccessful with non-coercive suggestion.

Commands resorted to only in life-and-death emergencies. Parent goes out of way to avoid coercion in his suggestions to child.

Scale No. 3.3 Accelerational Attempt (Acceleratory--Retardatory)

Rate the parent's striving to increase the rate at which the child's behavior is maturing. Does the parent deliberately train the child in various mental and motor skills which are not yet essential; or is the child left to "grow naturally"; or even shielded from accelerational influences?

This variable is restricted to purposeful teaching and training. It includes mental, motor, social, language, and personal skills. Disregard the effectiveness of the training. Consider the energy the parent exerts in striving to accelerate the child's behavior development.

Subjects child to regular and vigorous training to develop both essential and special skills.

Continually uses deliberate teaching to accelerate child in various skills, but with less than maximal vigor or regularity.

Frequently teaches and trains, but with restraint, and in limited number of skills.

Accelerational attempt is restrained and occasional only.

Almost entirely lacking in any deliberate training. Child left to grow "like Topsy."

Protects child from accelerating influences. Attempts to hold back rate of growing up.

Scale No. 4.1 General Babying (Over-helps--Refuses to help)

Rate the parent's tendency to help the child through the ordinary difficulties of every-day life. Does the parent insist on helping in situations where the child is quite capable; or does the parent withhold

aid even in major difficulties?

Rate relative to the child's ability. Disregard deliberate drill and training. This is a general variable including motor, mental, emotional, and social behavior. It applies only to tasks the child is attempting, not to parent-imposed requirements resisted by the child.

Continually helping child, even when child is fully capable and willing.

Usually helps more than needed. Seldom lets child struggle unsuccessfully.

Helps when needed, but not when child can get by alone.

Tends to withhold aid, letting child solve own minor problems.

Offers help after prolonged failure or in emergency.

Leaves child alone to solve even major problems, often refusing aid when requested.

Scale No. 4.2 General Protectiveness (Sheltering—Exposing)

Rate the parent's reaction to threats and hazards to the child's well-being. Does the parent tend to keep the child unnecessarily sheltered, and prevent difficulties from reaching the child? Or does the parent tend to expose the child to dangers, perplexities, and difficulties?

This is a broad variable, including protection from physical, bacterial, emotional, mental, and social hazards. Rate relative to the child's maturity. Disregard whether child is aware of protection. How much does the protective attitude of the parent tend to protect the child from experiencing difficulties?

Tends to shelter child from every imaginable slight discomfort or difficulty.

Not given to inventing imaginary hazards, but does protect from many trivial difficulties which child could handle.

Allows child to be exposed to many minor difficulties, but shelters from serious upsets even if purely temporary.

Lets child face own obstacles when there is no danger of lasting harm.

Exposes child to rather tough situations, unless danger is quite serious or situation acute.

Allows child to be exposed to major hazards, dangers, problems, suffering. Oblivious to hazards, or deliberately refrains from protecting child.

Scale No. 5.1 Readiness of Criticism (Critical—Uncritical)

Rate the parent's tendency to express an approval—disapproval attitude toward the child's behavior. Does the parent readily and vigorously express a reaction to things the child does; or is the parent non-committal, repressed, uninterested or stoical toward the child's actions?

Rate the tendency to express criticism regardless of whether it is approval or disapproval. Criticism may be verbal, gestural, or by facial expression or tone of voice—any signal to the child indicating approval or rejection of his behavior in a specific situation.

Alert to react to child's every move, regardless of how trivial.

Quick and free in expressing approval--disapproval. Criticism may be mild, or withheld, if the matter is obviously trivial.

Reacts freely when attention is called, but not alert for chance to criticise.

Tends to refrain from reacting critically in unimportant matters. Usually responds when pressed.

Withholds criticism unless greatly aroused. Tends to be poker-faced, noncommittal.

Gives no indication of either approval or rejection of child's acts, regardless of importance.

Scale No. 5.2 Direction of Criticism (Approval--Disapproval)

Rate the direction of the parent's critical reaction to the child's behavior. When the parent reacts does it tend to take the form of praise, approval, acclaim? Or does blame, disapproval predominate?

Rate only situations where a critical reaction occurs--a reaction to behavior, rather than general affectionateness or hostility. Criticism may be verbal, gestural, or by facial expression or tone of voice. It may be expressed either directly to the child or with the child as a witness. Rate relative to the merits of the behavior criticised.

Warm, unambiguous praise and commendation toward even rather ordinary behavior. Shortcomings overlooked or excused.

Emphasis on approval. Most disapproval is sugar-coated with simultaneous acclaim.

Balanced criticism. Praise, or disapprobation, predominates only as merited by child's behavior.

Tends to disapprove more readily than to approve. Most praise is tempered with faultfinding. Unduly critical of details.

Parent always finding fault. Ignores or belittles praiseworthy behavior, picking out minor details to criticise disproportionately.

Scale No. 6.1 Readiness of Explanation (Satisfies curiosity--Thwarts curiosity)

Rate the parent's tendency to satisfy the child's intellectual curiosity. Does the parent readily respond to the child's "Why?" and "How?" questions; or is the child thwarted in attempts to get information and explanation from the parent.

Disregard accuracy, depth, and honesty of answers. Note that explanations which are too ambitious, or too forced, may rate low; and the furnishing of specific references may rate high. Active coaching to think for self may rate high. Do not confuse with mental babying--a parent may refuse to do the child's simple thinking, and yet go out of his way to help with difficult explanations.

Never too busy to answer child as adequately as possible. Anticipates questions. Encourages curiosity with willing explanation.

Goes out of way to answer fairly involved questions, but sometimes postpones till child is older. May evade when very busy, or very tired.

Usually tries to satisfy child's curiosity. Sometimes loses patience with persistent "Why?"s.

Answers simple questions when in good humor and not preoccupied, but seldom goes beyond minimum needed to shut child up.

Explanations are grudging and reluctant where any mental effort is required. Parent often evades the issue.

Thwarts child's curiosity. Actively discourages questions with "Too busy", "You're too young to know", "Just because", etc.

Scale No. 7.1 Solicitousness for Child's Welfare (Anxious—Nonchalant)

Rate the parent's tendency to display over-concern for the child's well-being. Is the parent readily excited to overt anxiety all out of proportion to the importance of the situation? Or is the parent markedly calm, cool, and nonchalant, even in the face of critical danger to the child?

Consider the parent's net behavior regardless of the motives behind it. Include only behavior which is a potential stimulus to the child, impinging more or less directly upon his awareness. Include concern for both physical and mental health and comfort.

Given to severe, irrational anxiety on largely imaginary grounds. Readily panicked.

Chronic anxious tension over child, but more "jittery" than panicky. Given to "hunting for trouble."

Shows considerable anxiety when child is in any danger, but seldom loses rational control.

Somewhat solicitous, but minimizes hazards. Frequently shows concern, but without losing perspective.

Rarely worried or solicitous beyond needs of situation and responsibility as parent. Attitude more like that of teacher or nurse.

Nonchalant and seemingly unconcerned even in major matters. So unsolicitous as to appear neglectful or irresponsible.

Scale No. 7.2 Acceptance of Child (Devotion—Rejection)

Rate the parent's acceptance of the child into his own inner circle of loyalty and devotion. Does the parent act in such a way as to indicate that the child is considered an intimate and inseparable partner? Or does the parent act as though he resents the child's intrusion and rejects the child's bid for a place in his primary area of devotion?

Consider all evidence which in any way may impinge upon the child as acceptance-rejection, however subtle, vague, or indirect. It is not the parent's true feeling, but his attitude, as a functioning unit in the child's environment, which we are rating.

Parent's behavior toward child connotes utter devotion and acceptance into his innermost self, without stint or suggestion of holding back in any phase of his life.

Parent clearly accepts child. Includes child in family councils, trips, affection, even when it is difficult or represents considerable sacrifice.

A "Charter member" of the family, but "kept in his place". Parent accepts child in general, but excludes him from certain phases of parent's life.

Tacit acceptance. Excludes child so frequently that to the child

the rejection attitude may seem to predominate even though parent takes acceptance for granted.

Parent's predominant tendency is to avoid, repulse, and exclude the child, but without open rejection.

Child openly resented and rejected by parent. Never admitted to inner circle. Made to feel unwanted, ostracized.

Scale No. 8.1 Understanding (Keen--Obtuse)

Rate the parent's understanding of the child's abilities, needs, point of view, etc. Does the parent's behavior indicate a thorough and intelligent understanding of the child; or does it indicate a failure to appreciate the child's capacities and limitations, and an inability to meet the child on the child's own level?

This might be called "functional parental intelligence". It includes insight, foresight, child-empathy. Always rate in terms of the specific child.

Parent always sees subtleties of child's motivation; shows accurate appreciation of child's interests and degree of maturity.

Usually shows thorough understanding of child. Occasionally fails to see the point.

Has good grasp of every-day situations, but often misses the subtle angles.

Usually shows common sense where the point is obvious, but incapable of keen analysis.

Entirely lacking in subtlety; often misses the obvious.

Completely fails to see child's viewpoint, capacities, limitations. Expects entirely too much or too little. Fails to meet child on child's own ground.

Scale No. 8.2 Emotionality (Emotional--Objective)

Rate the emotionality of the parent's general behavior toward the child. Is the parent's reaction highly emotional; or is it consistently cool and objective?

Rate only in situation where there is sufficient cause for emotion to bring it out if it is there. Combine frequency and intensity of emotion. Combine direct expression of emotion and irrational distortion of policy due to emotion. "Emotion" as used includes manifestations of rage, panic, grief, disgust, love, mirth, or sympathy, where feeling predominates over reason.

Parent constantly giving vent to unbridled emotion in reaction to child's behavior.

Controlled largely by emotion rather than by reason in dealing with child.

Expression of emotion largely inhibited, but policy readily disorganized.

Emotion freely expressed, but actual policy seldom much disorganized.

Usually maintains calm, objective behavior toward child, even in face of strong stimuli.

Never shows any sign of emotional disorganization toward child, either directly or in policy.

Scale No. 8.3 Affectionateness (Affectionate--Hostile)

Rate the parent's expression of affection to the child personally. Does the parent manifest a warm, personal affection to the child; or a matter-of-fact, unemotional attitude; or definite antagonism?

Rate the attitude shown to the child, rather than the deeper one which affects the child only indirectly as through care, solicitude, or degree of devotion to the child's welfare.

Passionate, consuming, intense, ardent, uncontrolled.
Affectionate, warm, fondling, loving, expressive.
Temperate, fond, attached, forgiving, kind.
Objective, inhibited, neutral, matter-of-fact.
Cool, aloof, distant, forbidding.
Avoiding, annoyed, irritated, bothered.
Hostile, rejecting, disliking, blaming, icy.

Scale No. 8.4 Rapport between Parent and Child (Close rapport--Isolation)

Rate the closeness of the psychological relationship between parent and child. Do they show a high degree of rapport; or are they distant and out of touch with each other "spiritually", tending to be inhibited in each other's presence?

This variable includes mutual understanding, sympathy, confidence, and sharing of aspirations, intimate thoughts, and feelings. Rate it independently of the dominance-submission relationship. Do not confuse with antagonism-harmony.

Complete sharing of intimate thoughts and feelings. Implicit trust and confidence in each other.
Close mutual understanding and sympathy, but with occasional, temporary lapses.
Moderate degree of rapport in most situations; achieve close confidence in a good many respects, but fail in others.
Do not get along together any too well, but occasionally a close relationship is temporarily established.
Perfunctory relationship, superficial understanding, interest slight or forced. Tend to be inhibited in each other's presence.
Spiritually isolated. No sharing of confidence and aspirations. No active interest in each other.

List of Variables

The first six scales deal with the home as a whole. The others refer to a specific adult, usually the mother; but they can equally well be used to rate the father, grandmother, maid, or any other adult who has considerable contact with the child in the home.

For all variables the phrase preceding the colon describes the high end of the scale; the two extremes following the colon are in order, High score--Low score. No implications as to desirability are intended.

		Scale No.
1.	Adjustment of home: Well adjusted--Maladjusted	1.1
2.	Activeness of home: Active--Inactive	1.2
3.	Discord in home: Conflict--Harmony	1.5
4.	Sociability of family: Expansive--Reclusive	1.6a
5.	Coordination of household: Coordinated--Chaotic	1.7
6.	Child-centeredness of home: Child-centered--Child-subordinated	1.91
7.	Duration of contact with parent: Extensive contact--Brief contact	2.11
8.	Intensity of contact: Vigorous--Inert	2.12
9.	Restrictiveness of regulations: Restriction--Freedom	3.11
10.	Readiness of enforcement: Vigilant--Lax	3.12
11.	Severity of actual penalties: Severe--Mild	3.13
12.	Justification of policy: Rational--Arbitrary	3.14
13.	Democracy of policy: Democratic--Dictatorial	3.15
14.	Clarity of policy: Clear--Vague	3.16
15.	Effectiveness of policy: Successful--Unsuccessful	3.17
16.	Disciplinary friction: Contentious--Concordant	3.18
17.	Readiness of suggestion: Suggesting--Non-suggesting	3.21
18.	Coerciveness of suggestion: Mandatory--Optional	3.22
19.	Accelerational attempt (pushing): Acceleratory--Retardatory	3.3
20.	Babying: Over-helps--Withholds help	4.1
21.	Protectiveness: Sheltering--Exposing	4.2
22.	Readiness of criticism: Critical--Uncritical	5.1
23.	Favorableness of criticism: Approval--Disapproval	5.2
24.	Readiness of explanation: Satisfies curiosity--Thwarts curiosity	6.1
25.	Solicitousness for child's welfare: Anxious--Nonchalant	7.1
26.	Acceptance of child: Devotion--Rejection	7.2
27.	Understanding of child's problems: Keen--Obtuse	8.1
28.	Emotionality toward child: Emotional--Objective	8.2
29.	Affectionateness toward child: Affectionate--Hostile	8.3
30.	Rapport with child: Close rapport--Isolation	8.4

Instructions to the Rater

Making the Observations

1. Adequate pre-observation is, of course, essential for reliable ratings. Two two-hour home visits would seem to be a minimum for good results.

2. Familiarize yourself thoroughly with the scales before making the observations, so as to focus your attention on the variables to be rated.

3. Concentrate on building up an integrated impression of the ratee in terms of each variable. Concrete incidents and details must be subordinated to the generalized judgment which emerges from them.

4. Avoid direct quizzing of the ratee in terms of parent behavior. Center your approach around your interest in the child. You will learn a good deal about the parent indirectly through casual conversation and discussion of child problems. You are not primarily interested in the ratee's opinion of himself, but in what you can observe about his actual behavior toward the child.

5. Strive to cultivate a casual friendliness which puts the family at ease and encourages them to behave naturally rather than for "company" in your presence.

6. Avoid note-taking unless it proves definitely helpful in arriving at the thorough acquaintance with the family implied as a basis for the ratings.

Making the judgments

1. Your rating should represent your best estimate of the ratee's true position on the scale. Try to see beneath the special circumstances of the moment and judge the ratee's typical behavior for the current period.

2. Make use of all the evidence you have, including your total acquaintance with the ratee. But keep alert for significant changes or trends and always rate as of the present time rather than as of an average over a period of years.

3. Enough time should elapse between the observations and the rating so that the many specific events may merge into the 30 net judgments, and yet not so much that the judgments lose their sharpness. From two days to two weeks has been found to be a good range for this interval.

4. Rate up to ten cases in one set, completing all the ratings on a scale before passing to the next sheet.

5. Compare freely, one ratee with another, as the work on a given sheet proceeds. Work with a pencil, revising previous ratings as needed, so that when completed the sheet checks for absolute ratings, and for comparative rankings as well. (In doing several sets of ratings the rater may find it helpful when starting a new set of cases to enter in the margin of each sheet the scale positions of typical cases previously rated.)

6. When a home has received the usual period of observation it should be rated on all the variables on which the rater is able to make a judgment even if some of the ratings are recorded as "pure guessing" on the Certainty scale.

7. Consider every rating as applying to a specific child. Where there is more than one child under observation in a home each should have a complete set of parent ratings.

8. Sometimes it is impossible to make any judgment at all, and the rating is omitted. Some of the scales, for example, are inapplicable to the very young child. In such cases the reason for the omission should be written along the rating line.

9. Treat each scale as a smooth gradation from one extreme to the other. Use the cue points merely as points of reference in building up your concept of the total variable, rather than as discrete items to be checked. Good ratings will not tend to pile up at the cue points.

10. Keep your concept of each variable constantly refreshed by frequent study of the definitional material on the sheet. Avoid mere reference to the "name" of the variable; the name is merely a convenient handle for reference and may be very misleading if taken by itself to define the scale. Each variable is a complex of loosely correlated elements, and is defined by the descriptive paragraphs and all the cues on the sheet taken as a whole.

11. Throughout the process of observation and rating, and especially in making comparisons between ratees, it is essential that judgments of values be made on a purely objective or logical basis. Avoid all consideration of what is good or bad, wise or foolish, or optimal for ideal child training. Do not be concerned with what is average, above average, or below. Rate at the point which best fits the particular ratee.

The rating sheets

1. The entries in the boxes in the upper left part of the sheet are self-explanatory, and will be constant for a given set of thirty scale sheets. (The use of a gelatine duplicator will be found to facilitate the work of copying this material.)

2. The rating is made by placing an "X" on the rating line at the point which in your judgment best represents the ratee. Your rating may fall anywhere along the line from one extreme to the other, regardless of whether it falls opposite a cue point or somewhere between or beyond the cues.

3. The scales provide for five additional judgments to supplement each rating. These are intended to aid the rater in maintaining a uniformly careful technique, and to help evaluate the variability and importance of the several variables and the dependability of the ratings. In case there is no need for these supplementary data, nor concern that maximum care will be exercised in making the ratings, the supplementary judgments may be omitted entirely or in part.

The five supplementary ratings follow:

(a) Tolerance. Indicate by means of horizontal dashes (---) the limits on the scale between which you might place your "X" without being inconsistent with your judgment. In other words, place the dashes in such position that if another rater should place his "X" anywhere between them you would feel that he probably agrees with your judgment, whereas if his "X" falls outside your dashes his judgment probably differs significantly from yours. The interval between these points is called the "Tolerance."

(b) Range. The parent will vary somewhat up and down the scale during the period covered by the rating. Indicate this range by means of a convex curve (⌒) across the line at the highest point for which you have evidence and a

concave curve (⌣) for the lowest point. The interval is called the "Range."

The three remaining ratings are provided in the form of five-point numerical scales. The appropriate numbers are entered in the boxes immediately below the rating lines. These scales may be more fully defined as follows:

(c) Consistency. Within the period rated, from one situation to another, how much does the parent vary in his position on this scale?

(5) Ratee is incredibly consistent and predictable from one situation to another.

(4) Ratee deviates now and then from a clear-cut, typical point.

(3) A fairly definite central tendency, but with considerable variation within the rated period.

(2) Ratee quite inconsistent, but does show some central tendency.

(1) Ratee varies so much from time to time on this scale as to be almost entirely unpredictable.

(d) Saliency. If one is to understand the parent's total relationship to a particular child, certain characteristics stand out as essential while others seem unimportant. The term "Saliency" is here employed for the rater's estimated "weighting" of a rated variable in terms of its importance in the parent-child picture. A characteristic may be salient because it is frequently or intensely manifested, because the child is especially sensitive to it, because the rating itself is extreme, or because the variable is an intrinsically potent factor in child development. The salient ratings are those which should be high-lighted in a thumb-nail sketch of the ratee as a factor in the child's environment.

(5) One of the few most potent factors in this child's environment.

(4) Quite important in understanding the child; essential as part of the picture.

(3) Should probably be included in any adequate picture of the case.

(2) Of rather doubtful importance; would not alter the picture much if omitted.

(1) Plays no part in child's life; this rating could as well be ignored in understanding the child.

(e) Certainty. Your observations will have accumulated more evidence for some judgments than for others. How adequate do you feel your evidence to be for this rating? Would it take much new evidence to alter your judgment?

(5) Evidence is abundant and clear-cut for the occurrence of this characteristic at the rated value; it would require a great deal of new evidence to alter my judgment.

(4) Evidence is quite adequate; I feel reasonably safe in my rating.

(3) Evidence is rather incomplete, but I should not expect another home visit to alter my rating a great deal.

(2) Evidence is meagre; it wouldn't take much additional evidence to change my rating considerably.

(1) No tangible evidence; nothing more than a vague feeling which I can't justify.

It has been the usual practice to follow the main rating with the five supplementary ratings in the order given, before passing on to the next case on the sheet. But some raters have found it more helpful to reverse the order of the three graphic ratings, starting with Range, followed by Tolerance, and then the main rating.

4. Rater's "remarks" should be identified by the number of the child on the sheet. They should include such items as:

(a) Mention of anything unusual about the conditions of observation or rating.

(b) Exceptions taken to the definition or cues.

(c) Peculiarities of the ratee in regard to this variable, but not adequately covered by the rating.

(d) Any other qualifying remarks or mention of difficulties encountered in using the scale.

Scoring.

An ordinary millimeter scale may be used for scoring. Place the scale along the rating line so that the bottom end of the line falls just below the one-centimeter point, or at 9.5 millimeters. The rating may now be scored by reading the position of the "X" to the nearest millimeter, giving a possible range from 10 to 99.

In case single-digit scores are sufficiently accurate the second digit may be disregarded, giving a simple nine-point scale with centimeter units.

Values for Tolerance and for Range may be obtained, of course, by measuring the appropriate intervals to the nearest millimeter.

Scoring may be done quite rapidly by cutting off the millimeter stick at 100, sliding the lower end (at the scorer's left) along an appropriately placed heavy straight-edge with the left hand while the scores are read off and written in the space at the upper end of their respective rating lines.

SUMMARY

Starting with 30 variables of parent behavior which were developed in a previous study, this paper is concerned with devising a technique by which they may be measured. Parent-child behavior is too complex to be handled by the objective approach. The perceptual and integrative powers of the home visitor must be utilized, and provided with an instrument for their quantitative expression. The rating scale is ideal for this purpose if it can be made sufficiently accurate.

Several agruments are presented to indicate that the rating scale has greater potentialities than it has been given credit for, and several procedures are proposed for constructing optimally refined scales for the use of trained raters. A "graphic-parallel-vertical" format is described, employing "behavior cues" which have been scaled by a method based on the graphic judgments of experts. Certain criteria for good cues are formulated, and several supplementary ratings are proposed as a means of promoting more careful rating.

Finally, a battery of scales is presented for the use of the home visitor in appraising the child's environment in terms of 30 variables of parent behavior. A statistical report on reliability, validity, and norms will appear in a later paper.

REFERENCES

(1) Boring, E. G.: Is there a generalized psychometric function? Am. J. Psychol., 1924, 35, 75-78.

(2) Boring, E. G.: The logic of the normal law of error in mental measurement. Am. J. Psychol., 1920, 31, 1-33.

(3) Burtt, H. E.: Principles of Employment Psychology. Boston: Houghton Mifflin, 1926.

(4) Champney, H.: The Fels Parent-Behavior Rating Scales. Yellow Springs, Antioch Press, 1939.

(5) Champney, H.: Measurement of parent behavior as part of the child's environment. (Doctoral thesis on file at the Ohio State University Library.)

(6) Champney, H.: A statistical report on the Fels parent-behavior rating scales. (In preparation.)

(7) Champney, H.: The variables of parent behavior. J. Abnorm. and Soc. Psychol., (to appear soon).

(8) Champney, H. and Marshall, H.: Optimal refinement of the rating scale. J. Appl. Psychol., 1939, 23, 323-331.

(9) Conrad, H. S.: The California Behavior Inventory for Nursery School Children. Berkeley, Univ. Calif. Press, 1933.

(10) Conrad, H. S.: A statistical study of ratings on the California behavior inventory for nursery school children. Genet. Psychol. Monog., 1934, 16, No. 1.

(11) English, H. B., Welborn, E. L., and Killian, C. D.: Studies in substance memorization. J. Gen. Psychol., 1934, 11, 233-260.

(12) Freyd, M.: The graphic rating scale. J. Educ. Psychol., 1923, 14, 83-102.

(13) Goodenough, F. L.: Measuring behavior traits by means of repeated short samples. J. Juv. Res., 1928, 12, 230-235.

(14) Guilford, J. P.: Psychometric Methods. New York, McGraw-Hill, 1936.

(15) Macfarlane, J. W.: Studies in child guidance: I. Methodology of data collection and organization. Monog. Soc. Res. Child Develop., 1938, 3, No. 6.

(16) McGraw, M. B. and Weinbach, A. P.: Quantitative measures in studying development of behavior patterns (erect locomotion). Bull. Neurol. Inst. N. Y., 1936, 4, 563-572.

(17) Murphy, L. B.: Social Behavior and Child Personality. New York, Columbia Univ. Press, 1937.

(18) Olson, W. C.: The Measurement of Nervous Habits in Normal Children. Minneapolis, Univ. Minn. Press, 1929.

(19) Pratt, K. C., Nelson, A. K., and Sun, K. H.: The behavior of the newborn infant. Ohio State Univ. Stud., Contrib. Psychol., 1930, No. 10.

(20) Richards, T. W. and Powell, M. K.: The Fels child behavior scales. (To appear soon).

(21) Symonds, P. M.: Diagnosing Personality and Conduct. New York, Century, 1931.

(22) Symonds, P. M.: On the loss of reliability due to coarseness of the scale. J. Exper. Psychol., 1924, 7, 456-460.

(23) Thomas, D. S. and Associates: Some New Techniques for Studying Social Behavior. New York, Columbia Univ. Press, 1929.

(24) Thurstone, L. L. and Chave, E. J.: The Measurement of Attitude. Chicago, Univ. Chicago Press, 1929.

(25) Watson, D. L.: Scientists Are Human. London, Watts & Co., 1938.

THE BERKELEY GROWTH STUDY

HAROLD E. JONES AND NANCY BAYLEY[1]

The Berkeley Growth Study is one of several long-term developmental investigations comprised in the research program of the California Institute of Child Welfare.[2]

Specific aspects of this study have been reported in a series of articles and monographs listed in the attached bibliography; the present report deals with the general program, which has not been elsewhere described. The sixty-one cases originally included in this study were "normal" Berkeley children born in hospitals between September 25, 1928 and May 15, 1929. The families selected included only white, English-speaking parents who could be regarded as permanent residents of Berkeley and who were willing to cooperate in bringing their children to the Institute for the required series of examinations. In consequence of the selective factors involved, the group tends to be somewhat above the average in measures of socio-economic status (parental occupation, income, and education).

The original sample of 61 children (31 boys and 30 girls) has decreased slowly to 46 cases (24 boys and 22 girls) who remain available for regular examination at the present time (average age 12 years, as of January 1, 1941). Reduction in the sample has been due primarily to families moving out of town; so far as possible, relationships with these parents have been maintained through correspondence, with occasional opportunities to renew contact with the children. The original group has been augmented by ten younger cases, five girls and five boys, who have been followed through the same schedule, and now range in age from eight and one-half to ten years. Nine additional children (including a set of triplets) have been tested at less frequent ages on a less complete schedule.

The problems investigated have dealt chiefly with mental, motor, and physical development, observed from birth and with the plan of continuing to maturity. The program of psychological study is essentially descriptive, involving (a) the assessment, seriatim, of each individual's status in the group in a wide range of characteristics, (b) the study of trends and other age variations in status, and (c) the analysis of certain correlated factors. While interest has not been directed primarily toward personality nor toward attempts at the detailed interpretation of growth dynamics, considerable use has been made of qualitative observations and of material from interviews, in addition to the more largely quantitative records. The program of physical measurements and medical examinations is based on a schedule planned by Dr. Herbert R. Stolz. Dr. Bayley has been in direct charge

[1] *From the Institute of Child Welfare, University of California.*

[2] *With a varying organization of problems for the several groups of children, the samples studied have included, (a) children in attendance or in previous attendance at the nursery school, a cumulative sample started in 1927; (b) the Berkeley Survey, a representative group of children born in Berkeley during an eighteen month period in 1928-29; (c) the Berkeley Growth Study, begun in 1928; (d) twin studies, begun in 1928 and renewed in 1936; (e) the Guidance Study, undertaken with sub-samples of the Berkeley Survey in 1930; and (f) the Oakland Study of Adolescence begun in 1931-32.*

of the field program since 1931, and as will be noted in the appended list of publications, is the principal author of technical reports based upon the research to date.[3]

GENERAL PROCEDURES

As a first step in the study, cooperation was established with the attending physicians through the Hygiene Department of the University, and with the mothers at the hospital through the Institute of Child Welfare pediatrist. Within four days after birth a number of measurements were taken. These included tests of reflexes and of sensory functions, anthropometric measurements and records of blood pressure, pulse, respiration and temperature. During the first month the following additional information was secured: (a) The medical history of the mother, and pre-natal data, obtained by an Institute nurse in home visits and from physicians' and hospital records; (b) socio-economic data and family histories obtained from home visits by an Institute social worker.

Each child was brought to the Institute for a series of tests and observations at one-month intervals from birth to fifteen months, at three-month intervals from fifteen to thirty-six months, and at six-month intervals thereafter. Some additional visits have been made at irregular intervals for group observations or to collect material for special purposes. Below are listed the principal categories of data:

1. Mental tests

California First Year Mental Scale (26)	1 to 15 months
California Pre-school Mental Scale, Schedule I	18 to 60 months
Vocabulary tests (adapted from Thorndike CAVD)(30)	78 and 90 months
Dearborn Puzzle Board (30)	90 and 102 months
Stanford-Binet, 1916 Revision	72 and 84 months
Stanford-Binet, Form L	96, 108, and 132 months
Stanford-Binet, Form M	120 and 144 months

2. Motor abilities

California Infant Scale of Motor Development (36)	1 to 36 months
An unpublished continuation of the above motor scale	42 to 102 months
A series of tests of manual speed and dexterity	42 to 138 months
Foot-print records of standing and walking (21)	From age of first walking alone to 36 months.

[3]*Others assisting in the study have included Drs. Lotta V. Wolff, Ann Martin, C. C. Stevenson, H. K. Roe and O. G. Bates, pediatrists; Frances Welch and Jean Carter, social investigators; Agnes Covalt, research assistant in hygiene; Naye P. Nudge, Ruth N. Krause and Helene B. Schrader, statistical assistants. During the summer programs in 1938 and 1939, Dr. Pearl Bretnall, and for the academic year 1939-40, Dr. Mary K. Shirley, psychologists, have assisted in data collection. Dr. Herbert S. Conrad has served as statistical consultant, and since 1938 has been in charge of anthropometric examinations of boys. Dr. Nathan W. Shock has been in charge of physiological measurements since 1933.*

3. Reflex functions

Tests of approximately twenty reflex functions were made at each testing period through 36 months (24, 37).

4. Physiological measurements

Measurements of temperature (19), blood pressure (23), pulse rate and breathing rate were taken through 36 months. Starting at about 9 years, a series of physiological tests has been made on approximately two-thirds of the group. These include tests of exercise tolerance, basal metabolism determinations, and records of electrodermal reactions (galvanic skin reflex).

5. Anthropometric measurements

A schedule of 22 measures of bodily dimensions was taken monthly from one to 15 months; every 3 months to two years; semi-annually to four years; and annually to nine years (15, 16). After this age, in order to record changes related to the pubertal cycle, five of the measures have been recorded at six-month intervals, while the entire series, now including 27 measures, is administered once a year.

6. X-rays

X-rays of the chest and leg bones were taken at birth and at 1, 2, 3, and 4 years. With the addition of new equipment at the Institute, x-rays of the left hand and left knee were taken at 8 and 9 years, and semi-annually thereafter. These are being utilized for assessments of skeletal age by the Todd standards.

7. Photographs

Still photographs (nude) were taken at every second visit through 9 years, and at every visit thereafter. Motion pictures were made of crawling, creeping and early walking during the ages between six and eighteen months (39). Both still and motion pictures have been taken of the children in social play during picnics when they were near eleven and eleven and a half years.

8. Pediatric examinations

General physical examinations have been made by the Institute pediatrist annually beginning at two years. These include predominantly observations of conditions which might be of importance in the children's development—either physical conditions which themselves exhibit developmental changes, or health factors involving possible effects upon normal growth and development.

9. Emotional reactions and personality

In connection with all testing and measurement situations in infancy, records have been made of resistance and rapport, and of the duration, intensity and circumstances of crying when crying occurred (1).

Rating scales. Rating scales have been used consistently in connection with test situations, to record attitudes and incidental

behavior which might have a bearing on the interpretation of performance in these tests. As the children grew older these scales have been changed to evaluate more adequately behavior which is characteristic of the ages under observation. During the second year observations of behavior were made in standardized social situations, similar to those in the Marston and Berne series.

Projective Techniques. About two-thirds of the group, at 8.5 years, were observed in a creative art class, and samples of their paintings collected.

Other projective techniques introduced at this time (8.5 years) included a Rorschach test and a record of imaginative responses to a series of pictures similar to those employed by Murray. An interview schedule dealing with habitual interests and activities was also used. These procedures were repeated at 10.5 years. At 9.5 years selections from the Murray Thematic Apperception pictures were given; and were repeated at 11.5 years.

Home Interviews. Between the ages of nine and eleven years, Dr. Pearl Bretnall secured additional data on the personalities and home backgrounds of a small group selected for special study; and further case studies were made by Dr. Mary Shirley at eleven years.

Field Excursions. In order to have first-hand observations of social adjustments in free-play situations, the children have been invited to picnics or field excursions composed primarily of the members of the study. Several adults observed, and later rated the children on various aspects of their social behavior. Two such excursions have been held, to date, when the children were approximately eleven and eleven and a half years old.

In analyzing the data, the procedures have included a study of group trends in each field of development, an examination of interrelationships between the various fields through correlations at successive ages, and a study of each individual's growth curves considered separately and together. Standard scores have been used to facilitate comparisons of the data from different fields.

A further method of analysis which has been employed is the use of a simple mathematical function to describe approximately the growth of a given characteristic as measured in the individual child (20). The relative growth constant of this equation can be compared with similar constants for different characteristics of the child when fitted by the same equation, and with the constants for the same measurement in the other children. In the researches now in progress emphasis is being placed on studying interrelationships between various aspects of growth (mental, motor, physical and physiological), with attention also to emotional and environmental differentials. Although the individual curves are examined against the background of the whole group, the principal interest at the present time is less in the study of mass relationships than in the intra-individual comparison of growth curves for the various functions measured. During the past year a non-technical summary (7) was published by the University of California Press, dealing with home and family, early illnesses, growth in size, the trend of bone development, the development of some vegetative functions,

reflex patterns, the development of motor control, growth in intelligence, and the first ten years in review.

PUBLICATIONS

GENERAL

(1) Bayley, N.: A study of the crying of infants during mental and physical tests. J. Genet. Psychol., 1932, 40, 306-329.

(2) Bayley, N.: Some comparisons between growth in motor and in mental abilities in young children. Psychol. Bull., 1934, 31, 608. (Abstr.)

(3) Bayley, N.: A story of child care. San Francisco, Hearst Publications, 1937, pp. 26.

(4) Bayley, N.: The first Berkeley growth study. Berkeley, Univ. California Press, 1938, pp. 8.

(5) Bayley, N.: Mental and emotional growth in personality adjustment. Chapter II, Mental Hygiene in Modern Education. New York, Farrar & Rinehart, 1939, pp. 25-65.

(6) Bayley, N.: Mental and motor development from two to twelve years. Rev. Educ. Research, 1939, 9, 18-37 + 114-125.

(7) Bayley, N.: Studies in the development of young children. Berkeley, Univ. California Press, 1940, pp. 45.

(8) Jones, H. E.: Child study at the University of California. School and Soc., 1930, 31, 674-677.

(9) Jones, H. E.: The growth study as a psychological method. Psychol. Bull., 1935, 32, 538. (Abstr.)

(10) Jones, H. E. and Bayley, N.: Child Psychology. A home study course in the Extension Division of the University of California.

(11) Jones, M. C.: Neo-natal behavior. The Medical and Professional Women's Journal, 1933, (December), 362-364.

(12) Paraschivescu, O.: A comparative study of the socio-economic status of forty Berkeley families in 1929 and 1932. M.A. Thesis, University of California, 1933.

(13) Prentiss, S. W.: A preliminary study of data from a baby survey. M. A. Thesis, University of California, 1930.

(14) Schacht, H.: Watching intelligence grow. California Month., 1938, 41, 29-31 + 44.

PHYSICAL AND PHYSIOLOGICAL

(15) Bayley, N. and Davis, F. C.: Body growth during the first three years, and its relation to body build. Child Develop. Abstr., 1933, 7, 217-218.

(16) Bayley, N. and Davis, F. C.: Growth changes in bodily size and proportions during the first three years: A developmental study of sixty-one children by repeated measurements. Biometrika, 1935, 27, Parts I & II, 26-87.

(17) Bayley, N.: Some aspects of physical growth in young children. Psychol. Bull., 1935, 32, 526. (Abstr.)

(18) Bayley, N.: Growth changes in the cephalic index during the first five years of life. Human Biol., 1936, 8, 1-18.

(19) Bayley, N. and Stolz, H. R.: Maturational changes in rectal temperatures of sixty-one infants from one to thirty-six months. Child Develop., 1937, 8, 195-206.

(20) Jenss, R. M. and Bayley, N.: A mathematical method for studying the growth of a child. Human Biol., 1937, 9, 556-563.

(21) Rawlings, E. B.: Development of the longitudinal arch during the first two years of walking. M. A. Thesis, University of California, 1933.

(22) Wolff, E. and Stone, R. S.: Chest roentgenograms of non-tuberculous children suspected of being tuberculous. J. Am. Med. Assoc., 1930, 94, 458-460.

(23) Wolff, L. V.: Systolic blood pressure in early infancy. Arch. Pediat., 1930, 47, 165-170.

(24) Wolff, L. V.: The development of neuromuscular coordination during the first year of life. Transactions of the Second International Pediatric Congress. Acta Paediat., 1930, 9, 117-118. (Abstr.)

MENTAL DEVELOPMENT

(25) Bayley, N.: The consistency of mental growth during the first year. Psychol. Bull., 1931, 28, 225-226.

(26) Bayley, N.: The California first-year mental scale. University of California Syllabus Series, No. 243, 1933, pp. 24.

(27) Bayley, N.: Mental growth during the first three years. A developmental study of sixty-one children by repeated tests. Genet. Psychol. Monog., 1933, 14, 1-92.

(28) Bayley, N.: The maturation of mental functions. Chapter VIII, Readings in Psychology. New York, Farrar and Rinehart, 1935, pp. 219-251.

(29) Bayley, N.: The predictive value of several different measures of mental growth during the first nine years. Psychol. Bull., 1939, 36, 571-572. (Abstr.)

(30) Bayley, N.: Mental growth in young children. Part II, Thirty-Ninth Yearbook of the National Society for the Study of Education, 1940, Chapter II, 11-47.

(31) Bayley, N.: Factors influencing the growth of intelligence in young children. Part II, Thirty-Ninth Yearbook of the National Society for the Study of Education, 1940, Chapter III, 49-79.

(32) Bayley, N. and Jones, H. E.: Environmental correlates of mental and motor development: A cumulative study from infancy to six years. Child Develop., 1937, 8, 329-341.

(33) Jones, H. E. and Bayley, N.: Mental development and cultural-economic factors: A nine-year study. Psychol. Bull., 1938, 35, 681-682. (Abstr.)

(34) Jones, H. E. and Jorgensen, A. P.: Mental growth as related to nursery school attendance. Part II, Thirty-Ninth Yearbook of the National Society for the Study of Education, 1940, Chapter XII, 207-222.

MOTOR DEVELOPMENT

(35) Bayley, N.: The development of motor abilities during the first three years. Monog. Soc. Research Child Develop., 1935, 1, No. 1, pp. 25.

(36) Bayley, N.: The California infant scale of motor development. University of California Syllabus Series, No. 259, 1936, pp. 11.

(37) Wolff, L. V.: The response to plantar stimulation in infancy. Am. J. Dis. Children, 1930, 39, 1176-1185.

MOTION PICTURES

(38) Bayley, N. and Jones, H. E.: Case 75. I. 1 reel, 16 mm., silent. (The development of a child from one month through four years of age, as seen during testing procedures.)

(39) Jones, H. E. and Bayley, N.: The development of locomotion. 1 reel, 16 mm., silent. (Motion pictures of creeping, crawling and early walking of infants, illustrating developmental sequences, and the various forms of early locomotion.)

A COMPARISON OF ROUTINE HOSPITAL RECORDS OF BIRTH STATURE WITH MEASUREMENTS OF BIRTH STATURE OBTAINED FOR LONGITUDINAL RESEARCH

HOWARD V. MEREDITH AND JACK L. GOODMAN[1]

This study constitutes another unit of a pattern of research in which the first author has been interested for some years (8, 7, 4, 9). Succinctly, the interest is that of promoting increasingly fruitful collection of seriatim data by developing the requisite methodologic foundation.

Many of the so-called 'problems of analysis and interpretation of longitudinal data' can be shown to reduce to instances of insufficiently rigorous collection of data. In the scientific study of physical growth, perchance in the whole area of child development, this appears to be little appreciated and rarely emphasized. Consequently, studies focused upon improved collection of data are especially pertinent. It is not in the best interest of research effectiveness to collect data in a manner which unnecessarily burdens analysis and hampers decisiveness in the drawing of conclusions.

There are numerous indications that if studies on the physical status and growth of the individual child are to be efficiently executed and to yield sound knowledge there must be closer and more continuous attention to the collection of their basic data than many of those engaged in such researches now recognize. The accumulation of seriatim measurements of physical growth ought to presuppose well-considered reference to a substratum of information regarding (a) the relative precision of various anthropometric methods, and (b) the relation of measurement reliabilities to laboratory procedures and anticipated analyses. (See references 3, 4, 6, 7, 8, 9, 11.)

PURPOSE

The general objective of the present investigation was to add to the 'requisite base of technical information' discussed above. The specific aims were:

1. To determine the reliability of stature measurements taken on newborn infants under reasonably optimum laboratory conditions.
2. To compare routine hospital records of birth stature with records derived from the reliability study.
3. To evaluate the research usefulness of routine hospital records, i.e., their adequacy for studying the variability of birth stature, the position of individual newborns in the stature distribution, the variety of patterns of stature growth found in infancy, and so forth.

SUBJECTS

The subjects were one hundred white neonates, fifty males and fifty females. All were born at Hospital "X", a leading hospital in Iowa,

[1] *From Iowa Child Welfare Research Station, University of Iowa.*

between February 20 and April 20, 1941. They ranged in postnatal age from less than one hour to twenty-three hours.

Cases were taken serially as birth occurred except for the exclusion of infants considered "premature", infants placed under special observation, and infants physically deformed (e.g., talipes). In weight at birth the subjects varied from 2,580 grams to 4,730 grams, with a mean of 3,491 grams. The mean for the fifty males is 3,505 grams and that of the fifty females 3,476 grams.

STUDY OF RELIABILITY

The first objective in executing the plan of the present investigation was that of obtaining reliability findings for neonatal stature taken under relatively optimum laboratory conditions.

Anthropometric Technique

Instrument: The instrument used was a modification of the Baldwin Infant Measuring Board. The base of the board is one meter in length and 21 centimeters in width. At the head end of the base there is a fixed vertical plane, 15 centimeters in height. A vertical plane 7 centimeters in height slides freely in a brass groove extending along the lower two-thirds of the base. The board is inlaid with millimeter scales which lie parallel to the groove and are accurate to within one millimeter.

Technique: Each infant was placed on the board in dorsal recumbency. The anthropometrist then brought the subject's knees together and extended his lower extremities. With the aid of an assistant, the infant was lifted (in order to relieve any tension of the dorsal tissues produced through compression of the knees) and oriented in a position such that the long axis of the body paralleled the millimeter scales, the tragion-orbitale plane was at right angles to this axis, and the vertex made firm contact with the fixed upright of the instrument. The measurement was taken with the assistant maintaining the head position, with the anthropometrist using his right hand to keep the knees together and prevent their flexion, and with the anthropometrist employing his left hand to bring the sliding upright of the instrument firmly against the soles of the subject's feet.

Experimental Procedure

For any given subject, the procedure in amassing the reliability data was as follows:

1. The infant was brought from the nursery to a private examining room.
2. All clothing other than the standard navel dressing was removed.
3. Anthropometrist A placed the infant on the measuring board, carried through the technique described above, called off his stature reading to the assistant, and then left the room. The assistant recorded the reading.
4. Anthropometrist B entered the room and after first removing the infant from the board (an important safeguard in insuring an independent reading) duplicated the procedure followed by A.

B, of course, received no cue regarding A's reading.

In every instance J.L.G. served as anthropometrist B. H.V.M. served as anthropometrist A for about four-fifths of the cases and Dr. Virginia Knott or Miss Elizabeth C. Johnson for the remainder. All were trained anthropometrists having considerable experience in the collection of physical growth data on young children. The assistant varied from examination to examination. Those serving in this capacity were Dr. Knott, Miss Johnson, Miss Irene E. Harms, and H.V.M.

Reliability Findings

Each infant was employed as a subject for securing one pair of measurements. Consequently, the total reliability data consists of one hundred pairs of stature values. The mean age of the infants at the time of examination was nine hours, the range one to twenty-three hours.

The initial step in analysis was to discover whether the measurements of Anthropometrist B (the second series obtained) differed systematically from those of the first series. Subtraction of each value of the first series from its paired value in the second series resulted in the following tabulation:

Difference interval (mm.)	+6 to +9	+2 to +5	0 to ±1	-2 to -5	-6 to -9	-10 to -13
Frequency	3	24	48	20	3	2

This tabulation shows that the second measurements were higher than the first by two millimeters or more in 27 cases (27%), and lower than the first by two millimeters or more in 25 cases (25%). It may be concluded that the stature readings of Anthropometrist B register no tendency to differ systematically from those of Anthropometrists A.

When the above differences are grouped without regard to sign, they reduce to the following:

Difference Interval (millimeters)	Per cent
0 to 1	48
2 to 5	44
6 to 9	6
10 to 13	2

Specific findings supported by this distribution are:

1. Forty-eight per cent of the differences do not exceed one millimeter. In other words, for roughly one-half of the subjects examined, the stature obtained by one anthropometrist did not differ from that obtained by a second by more than 1/25 of an inch.
2. Seventy-seven per cent of the differences do not exceed three millimeters. That is, approximately 3/4 of the infants measured were found by two independent anthropometrists to yield stature values not disagreeing by more than 1/8 of an inch.
3. Differences not exceeding five millimeters, or 1/5 of an inch, are obtained for 92 per cent of the cases.
4. The mean difference is 2.44 millimeters, and the maximum

difference 12 millimeters.

In relation to mean stature at birth, the mean difference between anthropometrists is found to be less than one-half of one per cent. Relative to the average rate of growth during the first postnatal month, the mean difference between anthropometrists is equivalent to less than seven per cent.

Correlation of the one hundred stature measurements taken by Anthropometrist B with the one hundred taken by Anthropometrists A, using the Pearson product-moment method, gives a reliability coefficient of .973. This correlation is similar to those obtained by Knott (4) for stature of preschool children (r = .997) and by Marshall (7) for stature of elementary school children (r =.996).

COMPARISON OF EXPERIMENTAL RECORDS WITH HOSPITAL RECORDS

The writers derived a single stature record for each of the one hundred subjects examined. Wherever the difference between the two readings obtained for studying reliability did not exceed three millimeters, the record was taken as the mean of the two readings. In instances where the differences were greater than three millimeters, each anthropometrist made an additional determination and the record was taken as the mean of the four measurements.

The writers made their examinations without disturbing in any way the routine accumulation of stature records by the hospital staff. As a consequence, stature records for each of the one hundred neonates studied were available in the hospital files. These records were not copied from the files until after the writers had completed their own records.

Each stature record obtained by the writers was subtracted from the stature record for the same child carried in the hospital files. The following tabulation accrued:

Difference Interval (millimeters)	Per cent
+40 and up	2
+30 to +39	2
+20 to +29	9
+10 to +19	11
0 to ±9	54
-10 to -19	12
-20 to -29	7
-30 to -39	2
-40 and down	1

It will be seen that there is no systematic tendency for the one series of records to be either higher or lower than the other series. For twenty-four per cent of the subjects, the hospital record is larger than our record by 10 millimeters or more, while for 22 per cent of the subjects the hospital record is the smaller by 10 millimeters or more.

In view of the lack of a systematic difference between the two series of records, the following comparisons are markedly contrasting:

1. For 46 per cent of the cases, the difference between the hospital record and our record is 10 millimeters or more. Differences as large as 10 millimeters were obtained between stature records taken independently by two expert anthropometrists for only 2 per cent of the cases.
2. Forty per cent of the differences between the hospital records and our records are 12 millimeters or more, while only one difference between Anthropometrists A and B was as high as 12 millimeters. In contrast to no difference between Anthropometrists A and B exceeding one-half inch, differences between our records and the hospital records exceed one-half inch for upwards of 35 per cent of the cases and exceed one inch for approximately 15 per cent of the cases.
3. In relation to mean stature, whereas the mean difference between two expert anthropometrists was found to be less than one-half of one per cent (2.4 mm.), the mean difference between our records and the hospital records is more than four times this amount (11.2 mm.).

It will be appreciated that these comparisons imply even greater divergence than they express. This follows since the mean "error of measurement" of the writers' records should be considered as equal to approximately one-half of the difference between the measurements of Anthropometrists A and B.

The Pearson product-moment coefficient of correlation for the hospital records with our records is .741. It will be recalled that the reliability coefficient obtained on the same subjects was .973.

RESEARCH USEFULNESS OF HOSPITAL RECORDS FOR BIRTH STATURE

The present study was undertaken as a sector of research regarded as preliminary in initiating a study of individual patterns of growth covering the age period from birth to four years. The birth measurements for the projected study were to be obtained through Hospital "X", and the writers wished to know whether the hospital records for stature were reasonably accurate or whether it would be necessary to collect all the required data on the birth stature themselves. While the comparative findings already presented have clearly established the relative unreliability of the hospital records, it appears appropriate (especially since there is no reason to assume that the stature records of this hospital are taken any less rigorously than is typical of major hospitals where birth stature is obtained routinely) to consider their research usefulness in greater detail.

Would it be satisfactory to employ the files of Hospital "X" to obtain the mean birth stature of the one hundred newborns under study? The mean stature as obtained from the hospital records is 50.34 cm., and as obtained from our records 50.39 cm. The difference between these means (.05 cm.) is not statistically significant ($t = .17$). It can be concluded, assuming nothing beyond a measure of central tendency is required, that the routine hospital records are no less adequate than the records obtained specifically for research purposes.

Would it be satisfactory to employ the files of Hospital "X" to study variability in birth stature of the one hundred newborns under study? The standard deviation is 2.37 cm., as obtained from the hospital records and 1.60 cm., or one-third less, as obtained from our records. This difference is statistically significant ($F = 1.48$) at the five per cent level. The difference between the highest and lowest stature values of the hospital series is 12.0 cm., or 4 3/4 inches. Our range is 8.8 cm., or 3 1/2 inches - less than three-fourths of the hospital range. Such markedly divergent findings for the same one hundred subjects imply that the hospital records are decidedly too crude to be considered defensible for use in research on variability. It appears reasonable to infer that many of the published values for variability at birth may be spuriously high. In this connection it is not irrelevant to cite a text on infancy published as recently as 1930 (5). Here variability in stature at birth is treated in the single statement: "...babies have occasionally been observed to measure as much as thirty inches..." (5, p. 67) This figure stands a full eight and one-half inches above the maximum stature record in our series.

Would it be satisfactory to employ the files of Hospital "X" to determine the stature rank (position in the group) of individual infants? To secure information on this question, the one hundred subjects under study were arranged in rank order (a) according to the hospital records, (b) according to our records. Each individual was then assigned his percentile rank (a) in the hospital distribution and (b) in our distribution. The discrepancy between the individual's position in the two series was measured by computing the difference between his percentiles. The differences distribute themselves as follows:

Difference Interval	Per cent
0 to 9	43
10 to 19	23
20 to 29	20
30 to 39	9
40 to 49	3
50 to 59	1
60 to 69	1

Fifty-seven per cent of the subjects rank differently by ten percentile points or more in the hospital distribution than in our distribution. For one-third of the cases (34 per cent) the difference in rank exceeds twenty percentile points; for 14 per cent it exceeds thirty percentile points. Considered with reference to the reliability of the writers' records, these findings obviously preclude the use of the hospital records for obtaining the stature position of an individual infant in his group.

Would it be satisfactory to employ the files of Hospital "X" in studying problems of individual growth? The materials presented above clearly necessitate a negative reply. To many readers this will be in line with expectation. Nevertheless, the fact that such files cannot be used with reasonable confidence merits emphasis. In a longitudinal study reported in 1938 (10), hospital records for stature at birth were

correlated with measurements at three months of age obtained in a comprehensive research program. For 96 subjects the coefficient obtained was .54. In view of the fact that the correlation of our birth records with the birth records of Hospital "X" is as low as .74, it appears probable that the reported correlation of .54 is spuriously low. This probability is open to experimental evaluation. In other words, here is an instance where more careful attention to the collection of data would permit greater decisiveness in interpreting findings. Given a similar sample of 96 subjects for examination at birth and three months of age, it would be possible to obtain stature measurements known to be highly reliable and, from these, compute a coefficient justifying a reasonably precise statement regarding the true degree of relationship.

REFERENCES

(1) Bakwin, Harry, and Bakwin, Ruth Morris: Body build in infants. I. The technique of measuring the external dimensions of the body of infants. J. Clin. Investig., 1931, 10, 369-375.

(2) Cates, H. A., and Goodwin, J. C.: The twelve-day old baby. Human Biol., 1936, 8, 433-450.

(3) Davenport, C. B., Steggerda, Morris, and Drager, William: Critical examination of physical anthropometry on the living. Proc. Am. Acad. Arts & Sciences, 1934, 69, 265-284.

(4) Knott, Virginia B.: Physical measurement of young children: A study of anthropometric reliabilities for children three to six years of age. Accepted for publication in Univ. Iowa Stud., Stud. in Ch. Welf., 1941, 18, No. 3.

(5) Kugelmass, I. Newton: The story of infancy. New York, The Century Co., 1930. Pp. xvii, 327.

(6) Lincoln, Edward A.: The reliability of anthropometric measurements. Ped. Sem. and J. Genet. Psychol., 1930, 38, 445-450.

(7) Marshall, Everett L.: The objectivity of anthropometric measurements taken on eight- and nine-year-old white males. Child Develop., 1937, 8, 249-256.

(8) Meredith, Howard V.: The reliability of anthropometric measurements taken on eight- and nine-year-old white males. Child Develop., 1936, 7, 262-272.

(9) Redfield, Janet E., and Meredith, Howard V.: Changes in the stature and sitting height of preschool children in relation to rest in the recumbent position and activity following rest. Child Develop., 1938, 9, 293-302.

(10) Simmons, Catherine, and Todd, T. Wingate: Growth of well children: Analysis of stature and weight, three months to 13 years. Growth, 1938, 2, 93-134.

(11) Whitacre, Jessie: Standing heights of school children as determined by two techniques. Am. J. Phys. Anthropol., 1934, 18, 457-464.

FURTHER RESULTS ON THE TREND OF WEIGHT IN WHITE SCHOOL CHILDREN

GEORG WOLFF[1]

In an earlier paper on the trend of weight in white school children (1) of Hagerstown, Maryland, during the successive calendar years 1933 to 1936, the most striking result was the fact that the older school children, boys and girls from 12 years of age and over, exhibited a statistically significant increase of mean body weight. For the younger ages, on the other hand, there was no consistent change in mean body weight to be observed during the years of the period under examination. In addition, when these younger children, aged 6 through 11 years, are compared to the children of the same ages measured in the period 1921-27, no essential differences can be noted (2).

Measurements regarding height and weight have now become available for four more successive years, 1937-1940, in the same community, and therefore it is of special interest to examine whether or not the observed tendency in the pattern of growth, or more strictly speaking, in the trend of weight, is still continuing. This paper presents the pertinent data on weight and a similar analysis regarding height will follow in a subsequent publication.

Since methods of measurement and computation are exactly the same as described in the earlier paper, the details need not be repeated. However, the exact delineation of age is important in growth studies, and, therefore, it is well to mention that all the measurements have been taken as previously, each year within the first half of May. The age has been taken as of the last birthday before the date of examination, assumed to be May 1. Thus, the age designation 6, 7, or 8 years has the meaning of completed year of age. The children are grouped into age classes by single years. For example, the term 6 years of age covers the range from 6 to 7 years of age, or exactly from 6 years to 6 years and 364 days, the mid-point of the age group thus being 6½ years.

The observational data for the new four-year period, 1937 to 1940, covered in the present study, comprise a total of 19,997 weighings: 10,059 on boys and 9,938 on girls. Thus, the material dealt with considerably outnumbers the corresponding data for the previous four-year period, 1933 to 1936, covering a total of 14,401 weights: 7,134 on boys and 7,267 on girls (for more details cf. Table 1 of the earlier communication (1)).

The number of children weighed in each of the successive years is given in Table 1, which also contains the age range of the children.

[1]*Division of Public Health Methods, National Institute of Health, U. S. Public Health Service, Bethesda, Md., and the Milbank Memorial Fund.*

TABLE 1

NUMBER OF BOYS AND GIRLS WEIGHED IN HAGERSTOWN, MD., 1937 TO 1940

Year	Boys	Girls	Total	Age Range
1937	2,102	2,069	4,171	6-17 years
1938	2,372	2,398	4,770	6-17 years
1939	2,541	2,432	4,973	6-17 years
1940	3,044	3,039	6,083	6-20 years
All years	10,059	9,938	19,997	

The number of children examined increased from year to year, but it is quite obvious from the foregoing tabulation that there was an especially large increase in 1940, the last year of the present investigation. The reason for this is that in 1940, for the first time, all the boys and girls in the senior high school were included in the measurements. Therefore, the age range goes up to 20 years and the numbers of the older pupils from 15 years on are much increased as compared with the former calendar years. This may be seen still more distinctly from Table 2, in which the age distribution of the children examined is given by year of examination.

TABLE 2

NUMBER OF BOYS AND GIRLS WEIGHED IN HAGERSTOWN, MD., 1937 TO 1940, BY AGE

	Boys					Girls				
Age in years	1937	1938	1939	1940	1937-1940	1937	1938	1939	1940	1937-1940
6	129	126	172	130	557	113	155	168	155	591
7	182	235	237	235	889	200	207	216	278	901
8	220	219	254	246	939	240	231	203	253	927
9	256	227	248	266	997	240	251	242	255	988
10	253	277	246	251	1,027	252	264	258	245	1,019
11	253	265	272	270	1,060	250	250	266	238	1,004
12	279	258	283	282	1,102	285	276	255	275	1,091
13	266	286	281	303	1,136	256	306	286	267	1,115
14	167	260	316	276	1,019	171	264	303	279	1,017
15	70	153	167	298	688	51	153	166	297	667
16	20	53	50	216	339	11	35	52	221	319
17	7	13	15	151	186	--	6	17	170	193
18	-	-	-	79	79	--	-	-	88	88
19	-	-	-	27	27	--	-	-	18	18
20	-	-	-	14	14	--	-	-	-	-
Total	2,102	2,372	2,541	3,044	10,059	2,069	2,398	2,432	3,039	9,938

THE TREND OF WEIGHT IN WHITE HAGERSTOWN SCHOOL CHILDREN FROM 1933 TO 1940

The present records and computations enable us to extend the former growth study on the white school children of Hagerstown by single age years for a period of eight successive calendar years, i. e., from 1933 through 1940. Therefore, in order to obtain a clear view of the trend of weight over the whole period the weighted averages of both the four-year periods, 1933 to 1936 and 1937 to 1940, are added, and also the weighted average for the whole period. In addition, the adjusted values of Palmer's growth studies (2) covering a period of seven years in the preceding decade, from 1921 through 1927, are given in the respective summaries. Thus, the final comparison covers the years of two decades, 1921 to 1940, with a small interruption (1928-1932), and may serve as a good approach to delineate the growth trend in the weight of white school children in recent years.

The following tabulations, arranged as in the previous paper, contain for each age class the number of children examined in each calendar year, the range of the weight distribution, the mean weight in pounds and its standard error, the standard deviation, and the coefficient of variation as a relative parameter of dispersion.

BOYS AND GIRLS 6-7 YEARS OF AGE, FROM 1933 TO 1940

TABLE 3

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 6-7 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

	Boys					Girls				
Year	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	501		46.91 ± .273	6.11	13.02	529		45.26 ± .280	6.44	14.23
1937	129	30-69	46.88 ± .559	6.35	13.54	113	35-64	45.11 ± .481	5.11	11.34
1938	126	35-74	49.01 ± .613	6.88	14.09	155	30-64	45.98 ± .414	5.15	11.21
1939	172	35-64	47.79 ± .408	5.36	11.21	168	35-59	46.07 ± .388	5.03	10.93
1940	130	35-79	47.42 ± .583	6.65	14.02	155	35-69	46.95 ± .539	6.71	14.29
1937/40	557		47.72 ± .264	6.23	13.06	591		46.09 ± .227	5.52	11.98
1933/40	1058		47.34 ± .190	6.17	13.03	1120		45.70 ± .178	5.95	13.02
1921/27*	238		47.62 ± .357	5.51	11.57	237		46.57 ± .357	5.49	11.79

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

While it seemed, as explained in greater detail in the earlier paper, that there was in the four-year period 1933-1936 a slight decrease in average weight for both sexes, it does not hold true for the subsequent period 1937-1940. On the contrary, there was rather a slight increase of body weight in boys and girls although all the variations from year to year lie within the limits of chance fluctuation.

This can easily be verified by taking into consideration the standard errors of the means. However, the slight upward trend becomes somewhat more distinct if we consider instead of the values for the single years the weighted average of each of the four-year periods. For boys, the mean weight in 1933-1936 was found to be 46.91 ± 0.273 pounds, and in 1937-1940, 47.72 ± 0.264; thus the difference and its standard error amounts to 0.81 ± 0.378 pounds, or somewhat more than twice the standard error. For girls of the same age the corresponding mean value for each of the four-year periods was 45.26 ± 0.280 pounds and 46.09 ± 0.227 pounds. The difference is almost the same as that observed in boys, namely, 0.83 pounds, with a standard error of ± 0.360. This difference is 2.3 times its standard error, and can, as in the case of the boys, be considered significant from the statistical point of view. It appears then that for the first time since the start of the new seriatim examinations in 1933 there is also in the young school children an increase of weight that parallels the increase in the older ones. The drop in weight in 1935 and 1936 has been overcome in the following years. However, if we compare the average of the whole eight-year period 1933 to 1940 with that of the combined seven years of the period 1921-1927 there is rather a small decrease of body weight in the later decade for both sexes. It is not significant as can be easily seen from the tabulated values.

It is worth noticing that in every year of observation the boys exceed the girls in weight by 1 to 2 pounds. These consistent differences are beyond the limits of chance deviations and correspond to the sex differences in the biological pattern of growth.

The variability of weight, on the other hand, as measured by the standard deviation and the coefficient of variation does not exhibit any distinct tendency in this age group; in some years the parameters are higher for girls, in others higher for boys. It seems to hold true, however, that the variability constants in both sexes were somewhat larger in the period from 1933 on than in the previous period.

BOYS AND GIRLS, 7-8 YEARS OF AGE, FROM 1933 TO 1940

In this age group there appears again a slight increase of mean body weight for the period under observation. Although the fluctuations from year to year lie mostly within the limits of a single standard error, the general tendency to increase becomes more distinct when the averages of the periods, 1933 to 1936 and 1937 to 1940, are compared. For the boys these weighted averages amount to 50.89 ± 0.221 pounds and 51.99 ± 0.251 pounds, respectively. Thus the difference of 1.10 ± 0.334 pounds is statistically significant in the strict sense of the probability theory being 3.3 times its standard error. The picture is not quite so clear-cut in the girls of this age. There was a marked increase in weight in 1937, 1938, and 1939, when compared with the previous years of the total period, but in 1940 there was a set-back which amounts to almost 1 pound in mean body weight. Such a variation is still within reasonable limits of chance, if the mean values of the single calendar years and their relatively high standard errors are considered. The weighted averages in girls for the two four-year periods are 49.41 ±

TABLE 4

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 7-8 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	828		50.89 ± .221	6.37	12.52	842		49.41 ± .258	7.50	15.18
1937	182	25-74	51.37 ± .498	6.72	13.08	200	35-79	50.20 ± .531	7.50	14.95
1938	235	35-79	52.01 ± .455	6.97	13.40	207	30-99	50.08 ± .495	7.12	14.21
1939	237	35-89	52.56 ± .517	7.96	15.15	216	35-74	50.02 ± .418	6.14	12.28
1940	235	35-114	51.88 ± .524	8.04	15.49	278	30-79	49.26 ± .402	6.70	13.60
1937/40	889		51.99 ± .251	7.47	14.37	901		49.84 ± .228	6.84	13.72
1933/40	1717		51.46 ± 1.67	6.94	13.49	1743		49.63 ± .172	7.16	14.43
1921/27*	596		51.56 ± .270	6.60	12.80	573		49.96 ± .270	6.46	12.93

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

0.258 pounds (1933-1936) and 49.84 ± 0.228 pounds (1937-1940). Although there is a slight increase in the latter period, the difference of 0.43 ± 0.344 pounds has no statistical significance. Only the fact that this difference has the same direction as that for the boys may be noteworthy. When compared with the data collected in the former decade, the average of the period, 1933-1940, is still somewhat lower than in 1921-1927. The differences of 0.10 pounds in boys and 0.33 pounds in girls, of course, have no statistical meaning.

It may be seen also from the table that the boys surpass the girls in this age group by 1 to 2½ pounds in weight in every calendar year. The variability of weight presents no consistent tendency. Standard deviation and coefficient of variation tend, however, to rise somewhat when compared with the preceding age group (6-7 years), indicating the steadily dispersing effect of growth with increasing age. Otherwise, there is not visible any consistent similar trend with respect to age and sex. Both constants are higher in girls for the period 1933-1936, yet higher in boys for the period 1937-1940, as a whole. However, it is to be noticed again that for both sexes the variability constants are higher in the total period 1933 to 1940 than in the period of the preceding decade, 1921-1927.

BOYS AND GIRLS, 8-9 YEARS OF AGE, FROM 1933 TO 1940

The trend of body weight in this age is in almost perfect agreement with that of the preceding age group. The boys exhibit a distinct upward trend, the girls show hardly any noticeable change over the whole period of eight years. The weighted average in boys for the period 1933-1936 is 56.12 ± 0.245 pounds, for 1937-1940 it rises to 57.43 ± 0.287 pounds. The difference amounts to 1.31 ± 0.377 pounds and is

TABLE 5

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 8-9 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

	Boys					Girls				
Year	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	894		56.12 ± .245	7.32	13.04	891		55.19 ± .319	9.52	17.25
1937	220	35-99	57.03 ± .573	8.50	14.90	240	35-99	54.40 ± .533	8.25	15.17
1938	219	35-84	57.55 ± .569	8.42	14.64	231	40-89	55.40 ± .579	8.80	15.88
1939	254	35-89	57.01 ± .511	8.14	14.27	203	35-99	55.92 ± .615	8.76	15.67
1940	246	40-109	58.11 ± .638	10.01	17.23	253	40-84	56.04 ± .505	8.03	14.33
1937/40	939		57.43 ± .237	8.78	15.29	927		55.43 ± .277	8.44	15.23
1933/40	1833		56.79 ± .188	8.07	14.21	1813		55.31 ± .210	8.97	16.22
1921/27*	839		56.94 ± .262	7.60	13.35	811		55.14 ± .274	7.80	14.15

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

statistically significant. For girls the corresponding mean weights are 55.19 ± 0.319 and 55.43 ± 0.277 pounds, respectively, in the two periods. The small difference of 0.24 pounds has, of course, no statistical meaning, except that the same upward trend as in the preceding ages and as in boys may be noticed.

When the average of the whole period 1933-1940 is compared to the adjusted values of the findings reported by Palmer for the previous decade, there is no significant difference in weight. The boys have decreased by 0.15 pounds; the girls have increased by 0.17 pounds. Both differences are within the limits of a single standard error of the two averages, as can be seen from the table. Thus, the mean weights can be said to have been constant through the two decades, although there was for boys a significant increase of body weight for 1937-1940 when compared to 1933-1936.

In this age group the boys also show a greater body weight than the girls in each year of observation, although this sex difference is not quite so distinct as in the younger age groups. The variability of weight for the sexes is not consistent in the single years. The values of the standard deviation and coefficient of variation are larger and demonstrate a wider range of weight fluctuation for girls than for boys in the four-year period 1933-1936, but there is scarcely any difference between the sexes in the later period, 1937-1940, for which the constants are somewhat higher in boys. It is to be noticed again that the whole period of observation 1933-1940 tends to exhibit a greater variability of body weight than the period 1921-1927 of Palmer's observation. This holds true especially for girls as the tabulated parameters of variation show.

BOYS AND GIRLS, 9-10 YEARS OF AGE, FROM 1933 TO 1940

TABLE 6

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 9-10 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	934		61.65 ± .289	8.82	14.31	938		60.81 ± .397	12.15	19.98
1937	256	40-109	62.86 ± .557	8.91	14.17	240	40-104	61.40 ± .691	10.70	17.43
1938	227	40-134	63.14 ± .694	10.46	16.57	251	35-119	60.65 ± .648	10.26	16.92
1939	248	40-99	63.47 ± .602	9.48	14.94	242	40-104	61.90 ± .701	10.91	17.63
1940	266	40-109	63.97 ± .625	10.19	15.93	255	40-104	61.26 ± .606	9.68	15.80
1937/40	997		63.37 ± .309	9.75	15.39	988		61.30 ± .330	10.38	16.93
1933/40	1931		62.54 ± .212	9.30	14.87	1926		61.06 ± .256	11.24	18.41
1921/27*	978		62.59 ± .288	9.01	14.40	921		60.63 ± .317	9.63	15.88

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

Again an upward trend for boys is evident in this age group and it is particularly manifest when the weighted averages for the two four-year periods are compared. The mean body weight went up from 61.65 ± 0.289 pounds in 1933-1936 to 63.37 ± 0.309 pounds in 1937-1940. The difference of 1.72 ± 0.423 pounds is statistically significant in the strict sense of the probability theory, whereas the corresponding increase in girls from 60.81 ± 0.397 to 61.30 ± 0.330 pounds (0.51 ± 0.516 pounds) certainly lies within the limits of normal chance fluctuation. Thus far the behavior of the sexes is similar to that noted in the preceding two age classes (7-8 and 8-9); a significant increase of body weight has occurred in boys but in girls the increase is not statistically significant. Considering again the adjusted values of the period 1921-1927, the comparison with the later period 1933-1940 as a whole does not show any essential change. It means again that the average weight of the children, boys and girls, of this age has been fairly constant over the past two decades.

The variability of weight for all the different periods is larger for girls than for boys. It is again noticeable that in the present period of observations, as compared to the data of the former decade, the variability in body weight of girls has distinctly increased whereas the boys demonstrate rather stable values of the standard deviation and coefficient of variation throughout both decades. This finding confirms the results of Palmer's investigation. But it deserves to be mentioned that in the later four-year period of the present observation, 1937-1940, weight variability in girls also tends to decrease again and thus to approach the values found in the period 1921-1927 as well as those for the male school child. The values of the standard deviation for girls

are 12.15 in 1933-1936 and 10.38 in 1937-1940, as compared to 9.63 in 1921-1927. This could indicate that the years following the economic depression are again associated with a tendency toward diminishing group variability of body weight in girls similar to the trend in the years before the depression. The values for boys are consistently lower than for girls, as may be seen from the table.

BOYS AND GIRLS, 10-11 YEARS OF AGE, FROM 1933 TO 1940

TABLE 7

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 10-11 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	920		67.96 ± .343	10.40	15.30	934		67.41 ± .451	13.78	20.44
1937	253	45-104	68.07 ± .589	9.37	13.76	252	40-149	68.79 ± .997	15.35	22.31
1938	277	45-119	69.15 ± .581	9.67	13.99	264	45-124	68.92 ± .786	12.77	18.53
1939	246	45-119	69.51 ± .727	11.41	16.41	258	40-139	67.71 ± .793	12.73	18.80
1940	251	45-119	70.13 ± .759	12.02	17.14	245	45-154	69.81 ± .975	15.26	21.85
1937/40	1027		69.21 ± .330	10.59	15.30	1019		68.80 ± .439	14.00	20.35
1933/40	1947		68.62 ± .238	10.50	15.30	1953		68.14 ± .314	13.89	20.38
1921/27*	992		68.35 ± .344	10.84	15.86	925		67.49 ± .394	11.97	17.74

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

The upward trend of mean body weight is quite evident both for boys and girls of this age, although the trend was reversed for the male sex during the first four years of this period of observation. In the single years from 1933 to 1936 the mean weight of boys decreased by almost two pounds whereas that of the girls increased by the same amount. But when again comparing the weighted averages of the two four-year periods, 1933-1936 and 1937-1940, the tendency to increase becomes distinct. The mean body weight of the boys went up from 67.96 ± 0.343 to 69.21 ± 0.330 pounds, and that of the girls from 67.41 ± 0.451 to 68.80 ± 0.439 pounds. For the boys the difference between the two periods amounts to 1.25 ± 0.476 pounds, or 2.6 times, its standard error. The difference in the case of the girls is somewhat larger; 1.39 ± 0.629, but is not quite so significant due to the higher weight variability in girls and consequently the higher standard error. However, the latter difference, being 2.2 times its standard error, can be said to make the increase for both boys and girls statistically significant, in agreement with the general trend. It is hence noteworthy that the girls of this age, 10-11 years, indicate for the first time an appreciable increase of body weight, whereas the boys in all preceding age years exhibited such a tendency. On the average, the boys are still somewhat heavier than the girls although

there are already some exceptions in the single calendar years (1936, 1937), as can be seen from the tabulated values.

The mean weight for the whole period under observation, 1933-1940, exceeds somewhat that of the adjusted values for 1921-1927. However, the differences are small: 0.27 pounds for the boys and 0.65 pounds for the girls, and not yet significant. The variability of weight, as in the case of the younger age groups, is higher in girls than in boys throughout all the periods of observation, and it is again remarkable that in comparison with the period 1921-1927 only the girls show in 1933-1940 distinctly higher values of standard deviation and coefficient of variation. The corresponding values for boys are almost stable through all the periods.

BOYS AND GIRLS, 11-12 YEARS OF AGE, FROM 1933 TO 1940

TABLE 8

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 11-12 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	950		74.99 ± .403	12.41	16.55	971		76.03 ± .535	16.67	21.93
1937	253	45-169	75.09 ± .806	12.82	17.08	250	45-209	79.65 ±1.296	20.49	25.72
1938	265	50-149	75.52 ± .733	11.94	15.81	250	45-184	77.52 ±1.092	17.26	22.27
1939	272	55-134	76.86 ± .752	12.40	16.14	266	45-149	78.27 ± .990	16.15	20.63
1940	270	45-139	77.04 ± .852	14.01	18.18	238	50-144	77.50 ± .991	15.28	19.72
1937/40	1060		76.15 ± .399	12.80	16.81	1004		78.24 ± .546	17.30	22.11
1933/40	2010		75.60 ± .231	12.62	16.69	1975		77.15 ± .382	16.99	22.02
1922/27*	868		75.22 ± .430	12.66	16.83	798		75.90 ± .524	14.79	19.49

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

In this age group the upward trend of body weight is similar to that for the preceding age. It is especially distinct for girls, and in the strict sense, statistically significant if we compare the two four-year periods. The mean body weight for boys went up from 74.99 ± 0.403 pounds in 1933-1936 to 76.15 ± 0.393 pounds in 1937-1940; for girls, from 76.03 ± 0.535 pounds to 78.24 ± 0.546 pounds, respectively. Thus the increase in the girls: 2.21 ± 0.764 pounds, is for the first time in the present observation larger than in the boys, whose increase in weight from one period to another amounts in this age group to only 1.16 ± 0.563 pounds. It is further characteristic for this age group that the girls are heavier than the boys throughout the different periods compared. This was the case also for the combined years 1922-1927, as may be seen from the table. The adjusted values for this period are likewise somewhat lower than those of the combined years of

1933-1940, but the differences of 0.38 pounds in boys and also the larger one in girls of 1.25 pounds fall still within the limits of chance fluctuations. However, in contrast to the younger ages, for this age class the mean body weight cannot be said to have been constant over the last two decades. The tendency to increase is more obvious in girls, whose variability of body weight as measured by the variation parameters is likewise in contrast to the greater stability of weight of boys. This is clearly demonstrated by the tabulated values of the standard deviation and coefficient of variation, which are practically the same for boys in all periods, whereas the girls show again distinctly higher values in the period under observation, 1933-1940, than in the period of the former decade, 1922-1927.

BOYS AND GIRLS, 12-13 YEARS OF AGE, FROM 1933 TO 1940

TABLE 9

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 12-13 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	776		82.26 ± .537	14.97	18.20	804		85.15 ± .622	17.64	20.72
1937	279	50-159	83.36 ± .963	16.09	19.30	285	50-204	89.57 ± 1.174	19.82	22.12
1938	258	50-159	83.00 ± .869	13.97	16.82	276	45-239	91.71 ± 1.420	23.59	25.73
1939	283	50-174	85.01 ± .958	16.11	18.95	255	50-219	89.97 ± 1.261	20.14	22.39
1940	282	50-174	86.83 ± 1.058	17.77	20.46	275	50-164	89.12 ± 1.152	19.08	21.41
1937/40	1102		84.59 ± .483	16.03	18.95	1091		90.09 ± .694	20.66	22.93
1933/40	1878		83.63 ± .360	15.59	18.64	1895		87.99 ± .445	19.38	22.03
1922/27*	681		82.67 ± .595	15.52	18.77	614		86.13 ± .664	16.46	19.11

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

At this age girls approach puberty, but boys are not yet so far advanced in maturity. This different behavior is reflected also in the growth of the two sexes. The growth impulse is much stronger in the female sex, now beginning to "fill out the figure", and therefore the girls in this age of prepuberty show through all the periods combined and in every single year of the present series of observations a distinctly higher body weight than the boys. This was true also in the years 1933 and 1934 when a lower average weight was observed in comparison to previous and later years, a finding already pointed out in the preceding paper of this study (1). However, a general upward trend of body weight is evident for this age and statistically significant in both sexes. For boys the weighted average rose from 82.26 ± 0.537 pounds in 1933-1936 to 84.59 ± 0.483 pounds in 1937-1940; for girls from 85.15 ± 0.622 to 90.09 ± 0.694 pounds. Thus the respective

differences amount to 2.33 ± 0.722 pounds and 5.94 ± 0.932 pounds, the respective differences being 3.2 times its standard error in boys and 5.3 times its standard error in girls. Such an increase is definitely beyond chance fluctuations. Also, when the average of the whole period, 1933-1940, is compared to the average for 1922-1927 the increase of 0.96 pounds for boys and 1.86 pounds for girls is distinct and almost approaches statistical significance, especially in the female sex. This difference, of course, cannot be so large since the total period of 1933 to 1940 includes those years (1933 and 1934) which were associated with a definite drop of the mean body weight in both sexes. This transient decline was perhaps due to the economic crisis or its after effects. Whatever the reasons may be, it is none the less quite obvious from the present results that when we compare only the findings in the last four-year period, 1937-1940, with the corresponding figures of the previous decade, 1922-1927, the increase of mean body weight, 1.92 pounds for boys and 3.96 pounds for girls, becomes very distinct and is statistically significant.

The variability of body weight shows a trend similar to that described for the preceding ages. The value of the standard deviation (but not so much that of the coefficient of variation) rises in both sexes due to the steadily dispersing effect of growth. It is higher in girls than in boys. When compared with the 1922-1927 period only the girls exhibit an increased variability for 1937-1940 in particular, whereas the boys show in all periods almost constant values of the variation measures.

BOYS AND GIRLS, 13-14 YEARS OF AGE, FROM 1933 TO 1940

TABLE 10

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1933-1940 13-14 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1933/36†	667		92.88 ± .711	18.36	19.77	687		97.05 ± .697	18.27	18.83
1937	266	60-209	96.39 ± 1.206	19.67	20.41	256	60-199	99.96 ± 1.272	20.35	20.36
1938	286	60-179	95.00 ± 1.125	19.02	20.02	306	55-234	100.67 ± 1.189	20.80	20.66
1939	281	60-184	94.87 ± 1.057	17.72	18.67	286	50-184	102.06 ± 1.183	20.00	19.60
1940	303	60-194	97.57 ± 1.142	19.88	20.38	267	50-214	101.75 ± 1.271	20.78	20.42
1937/40	1136		95.98 ± .566	19.08	19.88	1115		101.12 ± .614	20.49	20.26
1933/40	1803		94.83 ± .443	18.81	19.84	1802		99.57 ± .463	19.64	19.72
1923/27*	466		92.57 ± .834	18.00	19.44	415		95.45 ± .862	17.56	18.40

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

The general trend in this age group of beginning puberty is exactly the same as discussed in the preceding age group. Both for boys and for girls the mean body weight went up considerably during the eight years of observation. Glancing at the tabulated values of the single calendar years, it can be seen that the maximum value for boys (in 1940) and for girls (in 1939) exceeds the minimum values (in 1934) for each sex by more than 10 pounds. This already indicates a significant increase in weight. However, since the standard errors of the means are high, especially in the critical years of 1933 and 1934 due to the relatively small sample, the weighted average of the two four-year periods, 1933-1936 and 1937-1940, will be compared. From 1933-1936 to 1937-1940 the weight of the boys increased from 92.88 $\pm$ 0.711 pounds to 95.98 $\pm$ 0.566 pounds while that of the girls increased from 97.05 $\pm$ 0.697 to 101.12 $\pm$ 0.614 pounds. The differences, for boys 3.10 $\pm$ 0.909 pounds and for girls 4.07 $\pm$ 0.929 pounds, are in both instances definitely significant. Almost as significant at this time are the differences between the averaged values of the two decades. The increase from the earlier decade, which comprises for this age the years 1923-1927, to the combined years 1933-1940 amounts to 2.26 $\pm$ 0.944 pounds in boys and 4.12 $\pm$ 0.978 pounds in girls. It would be still noticeably higher if out of the later decade only the last four years (1937-1940) were considered. Then the gain goes up to 3.41 and 5.67 pounds, respectively. This can be readily seen from the tabulated values. From a statistical viewpoint it would thus appear fairly certain that for the older school children who approach puberty there has been a definite increase in body weight in the two last decades, even if we allow for the marked drop in 1933 and 1934.

The feature described is more striking in girls at the age of prepuberty and beginning puberty than in boys. Is the conclusion then perhaps justified that concurrent with the evident acceleration of growth in body weight there has been an anticipation of sexual maturity during the last decades, as some authors maintain (for instance, see Peller (3))? Such an inference suggests itself, but would necessitate more data regarding the characteristics of sex maturation in girls, and, if available, in boys. Such data are unavailable at present but may be worth mentioning for a special study.

It can further be seen from the table that at this age the girls are considerably heavier than the boys. The differences vary between 3 and 5 pounds and are consistent in all the years of observation. On the other hand, the variability of body weight as measured by the standard deviation and the coefficient of variation is for the first time hardly greater for girls than for boys. However, in the period from 1933 the girls show again somewhat higher values of group variability than in the former decade.

BOYS AND GIRLS, 14-15 YEARS OF AGE, FROM 1934 TO 1940

Around this age boys enter the stage of adolescent acceleration in growth, which in general coincides with beginning of puberty; girls have almost completed growth and "filled out the figure". The tabulated values of the mean body weight makes it clear in the single years

TABLE 11

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1934-1940 14-15 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1934/36†	404		104.54 ± .944	18.97	18.15	458		104.38 ± .865	18.51	17.73
1937	167	60-169	106.00 ± 1.573	20.32	19.17	171	60-159	103.47 ± 1.258	16.45	15.89
1938	260	60-219	109.23 ± 1.368	22.06	20.19	264	70-194	108.86 ± 1.225	19.91	18.29
1939	316	65-199	109.57 ± 1.197	21.27	19.41	303	65-239	109.68 ± 1.156	20.13	18.35
1940	276	65-194	108.41 ± 1.297	21.55	19.87	279	55-229	111.14 ± 1.283	21.43	19.29
1937/40	1019		108.58 ± .670	21.39	19.71	1017		108.82 ± .621	19.81	18.20
1934/40	1423		107.43 ± .549	20.70	19.27	1475		107.44 ± .505	19.41	18.07
1924/27*	255		104.11 ± 1.108	17.70	17.00	220		104.02 ± 1.245	18.46	17.75

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

* Adjusted values from Palmer's investigations; compare footnote to Table 3 of the former paper (1).

of observation for this age there is scarcely any weight difference between the two sexes, apart from the year 1934, the first year of the present series of observations. The measurements for this age are available only from 1934 on and the reason has been explained in the earlier paper. The number of boys and girls examined in 1934 was small and thus does not count very much in the general trend of weight. But whatever may be the causes for the strikingly low body weight in boys in 1934 — effect of the economic depression or of some kind of selection — the upward trend is again distinct in both the sexes. Comparison of the weighted average for the periods 1934-1936 and 1937-1940 shows that body weight of boys rose from 104.54 ± 0.944 pounds to 108.58 ± 0.670 pounds; and of girls rose from 104.38 ± 0.865 to 108.82 ± 0.621 pounds. The mean weights are practically the same for boys and girls in each period. The increase in boys amounts to 4.04 ± 1.158 pounds, in girls to 4.44 ± 1.065 pounds, and is in both instances definitely significant. The increase from the period of the foregoing decade, 1924-1927, to 1934-1940 is also consistent in both sexes and significant; amounting to 3.32 ± 1.237 pounds and 3.42 ± 1.344 pounds, respectively. The gain would be still higher (4.47 pounds in boys and 4.80 pounds in girls) if out of the later decade only the period 1937-1940 is considered. Thus it may well be said that in this age of puberty, too, there is found a definite increase of body weight in the last two decades.

The variation parameters do not differ between the sexes very much at this age. However, in 1934-1940 the boys exhibit a higher variability than that noted in the period 1924-1927, while the girls demonstrate small change. After the latter have reached the age of puberty

they also show a greater stability of body weight. This is still more strongly indicated in the following ages in which the annual weight gain for girls is small, in contradistinction to boys, whose bony growth during puberty also conditions a further important increase of body weight.

BOYS AND GIRLS, 15-16 YEARS OF AGE, FROM 1935 TO 1940

TABLE 12

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1935-1940 15-16 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1935/36†	211		115.09 ± 1.351	19.63	17.06	180		107.45 ± 1.336	17.92	16.68
37	70	70-189	115.50 ± 2.475	20.71	17.93	51	80-149	106.72 ± 1.927	13.76	12.90
1938	153	65-184	118.97 ± 1.724	21.32	17.92	153	75-239	111.98 ± 1.534	18.98	16.95
1939	167	60-254	121.58 ± 1.810	23.40	19.24	166	70-219	116.17 ± 1.703	21.94	18.89
1940	298	70-209	120.60 ± 1.336	23.07	19.13	297	75-184	115.40 ± 1.056	18.20	15.77
1937/40	688		119.96 ± .859	22.52	18.77	667		114.14 ± .734	18.97	16.62
1935/40	899		118.82 ± .728	21.84	18.38	847		112.72 ± .644	18.75	16.63

† The values of the single years, 1933-1936, are given in the former paper (1) in greater detail.

For the first time since the beginning of the puberty stage the boys considerably surpass the girls in mean body weight. The girls have almost stabilized their growth at this age. This can be easily realized in the present investigation by noting the comparatively small increase of mean body weight as compared to the preceding age group. However, there is again in both sexes a distinct upward trend for the years under observation. Comparable data are available since 1935, but none for the former decade. For boys the weighted average rose from 115.09 ± 1.351 pounds in 1935-1936 to 119.96 ± 0.859 pounds in 1937-1940; for girls from 107.45 ± 1.336 pounds to 114.14 ± 0.734 pounds. Both differences, of 4.87 ± 1.601 and 6.69 ± 1.524 pounds respectively, are beyond the limits of chance fluctuation and mark a significant increase of weight during the last decade.

It is worth mentioning that at this age the boys show a considerably higher variability of body weight than the girls, indicating again the dispersing effect of growth, this time due to the greater vertical growth of the male adolescents. This pattern of sex differentiation in growth is still more emphasized in the following ages.

It may be further stated that from this age, 15-16 on, boys and girls examined in the calendar year of 1940 greatly outnumber those of the same age examined in former years. This is because in 1940, for the first time, the pupils of senior high school were included. For the same reason, two more ages are available in the present report in sufficient numbers for consideration, at least for the year 1940, as

the following data show.

BOYS AND GIRLS, 16-17 YEARS OF AGE, FROM 1936 TO 1940

TABLE 13

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1936-1940 16-17 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
1936	49	75-194	128.01 ± 3.276	22.93	17.91	33	75-149	112.65 ± 2.902	16.67	14.80
1937	20	70-274	131.75 ± 8.756	39.16	29.72	11	80-129	106.14 ± 3.753	12.45	11.73
1938	53	85-159	125.90 ± 2.330	16.96	13.47	35	80-154	113.50 ± 2.669	15.79	13.91
1939	50	75-214	129.20 ± 3.394	24.00	18.58	52	85-149	114.04 ± 1.874	13.52	11.85
1940	216	55-264	135.79 ± 1.736	25.52	18.79	221	85-229	119.69 ± 1.550	23.04	19.25
1937/40	339		133.03 ± 1.345	24.76	18.61	319		117.62 ± 1.138	20.33	17.28

Sex differentiation has proceeded at this age and so has the sex difference in mean body weight. Boys are heavier than girls by more than 15 pounds in almost all the years of observation, whereas in the preceding age group this difference averaged only 5 to 8 pounds. There is no doubt that such a difference is also statistically highly significant, as a glance at the tabulated values makes plain, especially in the year 1940 when the sample is large and the standard error therefore small.

The upward trend of mean body weight is still distinct in this shorter period of five successive years. It would not be satisfactory to compare the average of 1937 to 1940 with only one year, 1936; all the numbers in the single years are small except in 1940. It may, therefore, suffice to mention the upward trend of body weight similar to earlier ages and to record the weighted average of 1937-1940 only for completeness.

The variability trend is interesting insofar as the contrast in the development of the sexes in this age as compared to younger ages is still more manifest. Girls now show a smaller range of weight variation, so the variability constants are therefore distinctly higher in the male sex.

BOYS AND GIRLS, 17-18 YEARS OF AGE, FROM 1937 TO 1940, 17-18 AND UP IN 1940

There is visible also in this age group of adolescent males and females an upward trend of body weight during the last four years. Yet the numbers examined in 1937 to 1939 are much too small to verify the differences by a statistical test. It is noteworthy, however, that in 1940 when the numbers were adequate the girls of this age show exactly

TABLE 14

BIOMETRIC CONSTANTS OF WEIGHT (IN POUNDS) OF WHITE SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, 1937-1940 17-18, 18-19, 19-20, 20-21 YEARS OF AGE, AT MAY 1 OF EACH SUCCESSIVE YEAR

Year	Boys					Girls				
	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation	Number	Range	Mean ± Standard Error	Standard Deviation	Coefficient of Variation
17-18 years of age, at May 1										
1937	7	85-144	122.50 ± 6.701	17.73	14.47	—				
1938	13	75-214	136.35 ± 7.960	28.70	21.05	6	90-124	107.50 ± 3.727	9.13	8.49
1939	15	115-174	139.17 ± 3.902	15.11	10.86	17	90-134	112.20 ± 2.671	11.01	9.82
1940	151	90-194	139.92 ± 1.580	19.41	13.87	170	85-219	119.91 ± 1.368	17.84	14.88
1937/40	186		138.95 ± 1.441	19.65	14.14	193		118.85 ± 1.222	16.97	14.28
18-19 years of age, at May 1										
1940	79	105-219	143.13 ± 2.261	20.10	14.04	88	85-179	119.15 ± 2.237	20.98	17.61
19-20 years of age, at May 1										
1940	27	115-184	146.02 ± 3.730	19.38	13.27	18	80-194	121.67 ± 6.114	25.94	21.32
20-21 years of age, at May 1										
1940	14	105-219	154.64 ± 8.729	32.66	21.12					

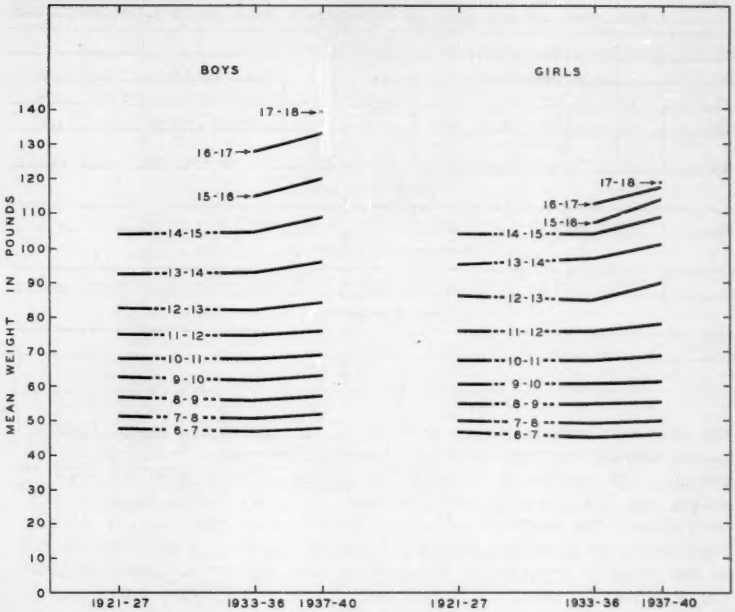
the same mean body weight as those of the preceding age group, 16-17 years, whereas the boys have again gained more than 4 pounds on the average. The pattern of variability is similar to that in the preceding age, but the variation parameters have diminished somewhat in both sexes. The standard deviation is again higher in boys, while the coefficient of variation due to the wide difference in mean body weight of the sexes is practically the same in boys and girls, though in 1940 the only year in which sufficient numbers were available, it is seen to be even somewhat higher in girls.

In the foregoing table are added the values obtained in the 1940 survey of the oldest students of the senior high school, 18-19, 19-20 years, and also 14 boys 20-21 years of age. No comment upon these figures can be made in the present study since comparable data for successive years are not yet available. So far as conclusions are admissible from these data, in view of the small numbers involved, it may only be mentioned that the male adolescents increase in body weight from year to year, whereas the young females show slight changes after the age year 16-17.

DISCUSSION

The secular upward trend in body weight of white school children appears still more distinct and conclusive in this investigation covering four calendar years 1937-1940 than it did in the investigation reported in the previous paper and which concerned the four years 1933-1936, including the period of the economic depression. The principal results for the whole period of eight successive calendar years, 1933-1940, are illustrated briefly in the following graph. A summary of the results

is further given in Table 15 which shows for all the age groups the gains (+) or losses (-) in mean body weight from (a) 1933-1936 to 1937-1940, and (b) 1921-1927 to 1933-1940.



The tendency toward an increase of mean body weight is present in practically every age group so far as the observations from 1933 through 1940 are concerned. From a statistical standpoint the increase is significant in young boys aged 6 through 9 years but in girls only 6-7 years of age. Then follows a short neutral age interval when the increase is practically the same for both sexes and relatively small in the eight successive years of observation. The tendency to increase becomes again more significant in the older school children, starting with the age 12-13 years in boys and 11-12 in girls, when the growth impulse becomes associated with sexual maturation. The increase is distinct in girls and stays so throughout the years of puberty to 15-16 years of age. Whereas the increase in young girls is smaller than in boys, the inverse is true from the age of prepuberty up to the years when girls stabilize their body weight. These older school girls have in the period of eight successive calendar years gained more than the

boys of the same ages. However, apart from these finer differences, the upward trend in mean body weight is evident in every age group.

TABLE 15

DIFFERENCE (INCREASE = +, DECREASE = -) IN MEAN BODY WEIGHT OF CHILDREN OF THE SAME AGE IN HAGERSTOWN, MARYLAND, DURING SUCCESSIVE PERIODS OF EXAMINATION

Age in years	Gain of mean body weight (pounds) in the periods from		Number of children examined			
	1933/36 to 1937/40 1/	1921/27 to 1933/40 1/	1933-36	1937-40	1921-27	1933-1940
		Boys				
6-7	+ 0.81 ± .378	- 0.28 ± .404	501	557	238	1058
7-8	+ 1.10 ± .334	- 0.10 ± .317	828	889	596	1717
8-9	+ 1.31 ± .377	- 0.15 ± .322	894	939	839	1833
9-10	+ 1.72 ± .423	- 0.05 ± .358	934	997	978	1931
10-11	+ 1.25 ± .476	- 0.27 ± .418	920	1027	992	1947
11-12	+ 1.16 ± .563	+ 0.38 ± .314	950	1060	868 2/	2010
12-13	+ 2.33 ± .722	+ 0.96 ± .695	776	1102	681 3/	1878
13-14	+ 3.10 ± .909	+ 2.26 ± .944	667	1136	466 4/	1803
14-15	+ 4.04 ± 1.158	+ 3.32 ± 1.237	404 5/	1019	255 6/	1423
15-16	+ 4.87 ± 1.601	——	211 7/	688		
16-17	+ 5.02 ± 3.541	——	49 7/	339		
		Girls				
6-7	+ 0.83 ± .360	- 0.87 ± .399	529	591	237	1120
7-8	+ 0.43 ± .344	- 0.33 ± .320	842	901	573	1743
8-9	+ 0.24 ± .422	+ 0.17 ± .345	891	927	811	1818
9-10	+ 0.51 ± .516	+ 0.43 ± .407	938	988	921	1926
10-11	+ 1.39 ± .629	+ 0.65 ± .504	934	1019	925	1953
11-12	+ 2.21 ± .764	+ 1.25 ± .648	971	1004	798 2/	1975
12-13	+ 4.94 ± .932	+ 1.86 ± .799	804	1091	614 3/	1895
13-14	+ 4.07 ± .929	+ 4.12 ± .978	687	1115	415 4/	1802
14-15	+ 4.44 ± 1.065	+ 3.42 ± 1.344	458 5/	1017	220 6/	1475
15-16	+ 6.69 ± 1.524	——	180 7/	667		
16-17	+ 4.97 ± 3.117	——	33 7/	319		

1/ See the following footnote regarding the varying periods of observation for the children 11-12 years of age and over.

2/ The period comprises the calendar years 1922-27; 3/ 1923-27; 4/ 1924-27 only.

5/ The period comprises the calendar years 1934-36; 6/ 1935-36; 7/ 1936 only.

Thus far the findings of the former study are corroborated and in some respects enlarged and corrected by the present data. Some small variations which seemed in the youngest age group (6-7 years) to indicate a certain, though not statistically significant decrease in mean body weight from year to year are smoothed out when the four years of the former observation and the four years of the present one are combined and the weighted averages are compared. Likewise, some of the apparent inverse trends in boys and girls aged 9-10 and 10-11 years, which were mentioned in the earlier study, are beginning to disappear. Of course, the most outstanding finding of the former report, the phenomenon of a conspicuous increase in body weight of the older school children in Hagerstown, is still more strongly emphasized by the present results. The drop in body weight observed in the years 1933 and 1934 concerned particularly these older boys and girls, but this decline was followed soon by a consistent rise through all the subsequent years. The phenomenon becomes more clearly marked in comparing the two four-year periods and the increase for the older school children also holds when the means for 1933-1940 are compared with the combined values of Palmer's earlier study in the same community for the period 1921-1927.

Therefore, here is a real increase in body weight of older children during the last two decades whereas the average weight of the younger children, 6 to 11 years of age, has remained fairly constant during this time.

There is another noteworthy point relating to the pattern of growth and it concerns the variability of body weight. In both periods of the present observation, 1933-1936 and 1937-1940, the girls exhibit consistently through all the available age groups distinctly higher variability constants in comparison with the earlier period, 1921-1927; in contradistinction, the boys show through all the different periods of observation rather constant values of the standard deviation and coefficient of variation, except in the highest age groups. It might be reasonable to assume that the greater variability of girls in recent periods of investigation is more or less caused by the greater gains of weight in these later years and parallels thus the greater variability in boys at the beginning of adolescence when a new growth impulse comes into power. From this point of view, the greater variability of body weight at different times of examination or at different stages of growth development of the sexes is no more than the expression of a physiologically widening growth range, i. e., the dispersing effect of growth.

The question suggests itself: What are the possible reasons for the distinct increase of mean body weight in the white school children in Hagerstown? The phenomenon of a general increase in growth of children is world wide and has been reported from many communities, just recently from Zuerich, Switzerland, in the careful investigation (4) of the Statistisches Amt der Stadt Zuerich. The cited study covers around 30,000 measurements of school children, boys and girls, over the years 1934 through 1939. The period under observation is almost the same as in our present investigation, and the results regarding the general increase of mean body weight and height are similar. However, the reasons for such a universal tendency are not yet clear enough and will be discussed in greater detail in a later paper, in which the height constants are calculated for the same generations of children so far as they are available, and some relations between height and weight will also be presented. For the present, it will be recalled that body weight in particular is a highly variable element in growth and probably depends much more than growth in stature on environmental factors, i. e., the influence of varying natural and social conditions. But climate and other elements of the natural surroundings have hardly changed in the short period of 7 or 8 years and a starvation period also is out of the question in the United States or in Switzerland for the period under observation; in the contrast to the experience in Germany and Austria during and after the war of 1914-1918 (5). Thus, for a possible explanation it may well be necessary to have recourse to a consideration of recent changes of a general character that have taken place within populations such as the general birth decrease which in recent times is common in almost all countries of western civilization. It is certain that with the decline of birth rate and, consequently, with the diminution in size of family, the mean income per head has increased. This may have something to do with the general tendency toward an increase

of growth since it is a well established fact as shown in various growth studies of children of different social strata (6, 7, 8) that the general improvement of social and economic conditions, especially nutritional conditions, shows its influence also upon bodily dimensions during the years of growth.

In this connection, it must be mentioned, as recently pointed out by Ciocco (9) in a special population study in Washington County, Maryland, from 1898 to 1938, that the highest socio-economic class has essentially increased its contribution to the natality of the population at large. Yet on the other hand, as the author explains in a further study with regard to parity order of birth and socio-economic status (10) the diminution in size of family concerns in the first rank the same upper social strata. For the highest socio-economic class (Class I: owners, managers, officials, and professional men) he found that "the reproductive pattern in recent years has been characterized by a true limitation of the size of the family." For the lower groups (Class II: semi-professional and skilled workers; Class III: semi-skilled and unskilled laborers) which of course furnish the bulk of the births, the parity pattern of the mothers has not essentially changed; hence "the principal feature of the change in the reproductive pattern of Class II and III appears to be a diminution in the number of women who reproduce at all." Thus it can be more readily explained that in the face of a declining birth rate of the population as a whole the highest socio-economic class has significantly increased its proportional contribution to the natality of the population, especially relative to births of parity order 1 to 3 as demonstrated by Ciocco's new figures. Whatever the outcome of this differentiation in class fertility may be with regard to the final size of the familial aggregate, the change of this reproductive pattern offers a possible explanation of the striking growth increase in school children during the last decades. Since children of the higher socio-economic level are on the average heavier than others, an increase in the proportion of such children in a population would conceivably raise the level of the average weight of the children of the population. It would be worth, anyway, a closer examination to study the trend of growth by parity order of birth on the one hand and to consider the socio-economic status of the parents at the same time.

SUMMARY

This report presents the second of a new series of studies dealing with the pattern of growth and annual variations in growth constants of white school children of Hagerstown, Maryland, The material was collected by the United States Public Health Service during the years from 1937 through 1940 and is the direct continuation of the earlier growth study in the same community for the years from 1933 to 1936. Thus data on a period of eight successive calendar years are available in the present paper, and they are further compared with Palmer's previous studies on mean body weight of white school children of the same community during the preceding decade.

The observational data of the four-year period, 1937-1940, comprise nearly 20,000 weight measurements, 10,059 taken on boys and 9,938 on

girls, aged 6 to 17 years during 1937-1939, and 6 to 20 years in 1940. Together with the data on 14,401 weight measurements taken in the preceding four-year period, 1933-1936, the observations used in the eight successive calendar years cover a grand total of almost exactly 34,400 measurements, which were always taken during the first days of May of the respective years. In addition to the mean weight constants for the single calendar years, there are given here the weighted averages of each four-year period, 1933-1936 and 1937-1940, to be compared with each other; and the average for the total eight-year period to be compared with that of Palmer's investigations covering the combined years 1921-1927.

On the whole, the findings reported here tend to increase the validity of the previous observations. They may be summarized as follows:

1) There is a distinct secular upward trend in weight for all age groups and for both sexes which can be recognized when the two four-year periods 1933-1936 and 1937-1940 are compared. In boys the mean increase of body weight during the total eight years of observation is visible in every age group, though not to the same degree through the entire period of development, yet it is statistically significant in the ages from 6 years on (6-7 years) to 15 years (15-16 years). In girls it is beyond the limits of mere chance fluctuations in the lowest age group, 6-7 years, hardly so in the following three age groups from 7 years of age (7-8) through 9 years (9-10), but from 10 years on for the successive years of observation the weight increase again becomes distinct and statistically significant. These final results complete and correct, in some minor points of the trend, the earlier findings for the first four years, 1933-1936, when a statistically significant increase was found only in the older school children, boys and girls, from 12 years of age and up.

2) When the combined values of the total period 1933-1940 are compared with the values obtained in the previous decade, for the combined years 1921-1927, there is no essential variation in body weight of the young school children from 6 years up to 11 years of age. For children of these ages the body weight remained fairly constant over the last two decades, with some minor fluctuations within the limits of chance. The first noticeable increase occurs for boys of 12-13 years of age, and for girls, 11-12 years old, and becomes more striking for the subsequent ages of beginning puberty. For the last age group, 14-15 years, for which data are available from the preceding decade, the increase in mean body weight from 1924-1927 to 1933-1940 amounts to 2.32 ± 1.237 pounds for boys and 3.42 ± 1.344 pounds for girls. The gain is still distinctly higher if out of the later decade only the last four years (1937-1940) are considered.

3) The variability parameters: standard deviation and coefficient of variation do not show any consistent secular trend in the single calendar years of the present observation for children of the same age. On the other hand, as is well known, with increasing age the values go up due to the dispersing effect of growth. As regards the sex difference in variability of body weight, there is hardly any striking difference in the values obtained for the younger age groups from 6 to 8 years.

However, from the following age group on, 9-10 years up to 13-14 years, the girls demonstrate a higher variability in almost all periods of observation. The greater stability of body weight in boys disappears in the next age group, 14-15 years, when the new growth impulse in the male becomes effective. From this age on the variability parameters are mostly higher in boys, indicating as a rule the wider range of their body weight as compared to girls of the same chronological age.

4) The phenomenon of a real increase in body weight of the white school children in Hagerstown over the last decade, particularly of the older children over the last two decades, appears incontestable. What the reasons are can not be easily grasped. The phenomenon is world-wide and therefore we could assume that the general increase of bodily dimensions may be indirectly associated with other world-wide phenomena such as the birth decrease in the western countries and the connected diminution in size of family. This striking socio-biological phenomenon of recent times has, of course, increased the mean income per head, whether real wages have or have not risen. The assumption of a causative connection is suggestive after some very reliable investigations have revealed a distinct association between growth of children and well-being of the parents. Such a bearing of the general birth decrease upon children's growth appears the more reasonable after it has become evident from new figures of the population trends in Washington County that the highest socio-economic class has, since about 1920, steadily increased its proportional contribution to the natality of the population.

REFERENCES

(1) Wolff, Georg: A study on the trend of weight in white school children from 1933 to 1936. Material based on the examinations of pupils of the elementary schools in Hagerstown, Maryland. Child Develop., 1940, 11, 159-180.

(2) Palmer, Carroll E.: Variation of growth in weight of elementary school children, 1921-1928. Pub. Health Rep., 1933, 48, 993-1005. (Reprint No. 1591)

Idem: Growth and the economic depression. A study of the weight of elementary school children in 1921-1927 and in 1933. Pub. Health Rep., 1933, 48, 1277-1292. (Reprint No. 1599)

(3) Peller, Sigismund: Growth, heredity and environment. Growth, 1940, 4, 277-289.

(4) Statistisches Amt der Stadt Zürich: Körpermessungen an Zürcher Volksschülern. Zürcher Statistische Nachrichten, 1940, 17, 181-240.

(5) Wolff, Georg: Die Nachwirkung der Kriegshungerperiode auf das Schulkinderwachstum. Leipzig, Johann Ambrosius Barth, 1932.

Idem: Increased bodily growth of school children since the War. Lancet, April 27, 1935, 1006.

(6) Oettinger, W.: Anthropometrische Untersuchungen an Breslauer und Charlottensburger Schülern. Ztschr. f. Hyg. u. Infektionskr., 1922, 98, 338-414.

(7) Palmer, Carroll E.: Growth and the economic depression. l.c.

Idem: Further studies on growth and the economic

depression. Pub. Health Rep., 1934, 49, 1453-1469. (Reprint No. 1660)

(8) Jenss, Rachel M.: Gain in weight and its association with ancestry and economic status. Human Biol., 1940, 12, 532-544.

(9) Ciocco, Antonio: The trend in the proportional contribution of the socio-economic groups to natality. A report based on the births in Washington County, Md. from 1898 to 1938. Human Biol., 1940, 12, 188-202.

(10) Idem: Parity order of birth, age of mother, and socio-economic status. Human Biol., 1941, 13, 64-87.

STUDIES ON DENTAL CARIES
XI. SEX DIFFERENCES IN THE CARIES SUSCEPTIBILITY OF THE VARIOUS MORPHOLOGICAL TYPES OF PERMANENT TEETH[1]

HENRY KLEIN[2] AND CARROLL E. PALMER[3]

In a recent study (1) it was shown that, at the same chronological age, girls have a larger total number of permanent teeth with evidence of caries experience than boys. Teeth with evidence of caries experience include those which are decayed, indicated for extraction, extracted, or filled. These are designated "DMF teeth" (2).

Observation of the numbers of erupted teeth in the two sexes shows that, at the same chronological age, girls also have a larger total number of erupted permanent teeth than boys. This latter finding arises out of the fact that girls erupt their permanent teeth at earlier chronological ages than do boys (3). At the same chronolcgical age, therefore, the permanent teeth of girls have been exposed in the mouth (post-eruptively) to the risk of attack by caries a longer duration of time than those of boys (4).

The sex difference in numbers of DMF teeth noted above is explained by this sex difference in posteruptive tooth age. Thus, when correction is made for the tooth age difference between the two sexes it is found that both sexes show approximately the same total number of DMF permanent teeth. From this finding the conclusion has been reached that there exists (at least up to chronological age 15 years) no significant sex difference in caries susceptibility (1). These findings and conclusions were based on a study of all the permanent teeth taken as a whole in a group of approximately 4,400 elementary school children.

It is the purpose of the present paper to consider whether or not a sex difference exists in the caries susceptibility of the individual morphological types of permanent teeth. For this purpose the material on the 4,400 grade school children is used, combined with additional observations on a group of approximately 2,000 high school children. Accordingly, the findings given here are based on the individual types of permanent teeth in a total of 6,257 girls and boys attending the grade and high schools of Hagerstown, Maryland, and environs.

FINDINGS

The age and sex distributions of the children examined are shown in Table 1. The numbers of DMF teeth of particular morphological types found in these children are given in Tables 2 and 3. The DMF values shown are averages for the single types of teeth of either the right or left sides of the mouth. The DMF values are given in this form since the caries susceptibility of the same type of tooth on each side of the mouth is about equal. It may be noted that in general the DMF values for each tooth are higher in the girls than in the boys at the same chronological age.

[1] *From the Division of Public Health Methods, National Institute of Health.*
[2] *Dental Officer, U. S. Public Health Service.*
[3] *Passed Assistant Surgeon, U. S. Public Health Service.*

Figure 1
Maxillary 1st premolars

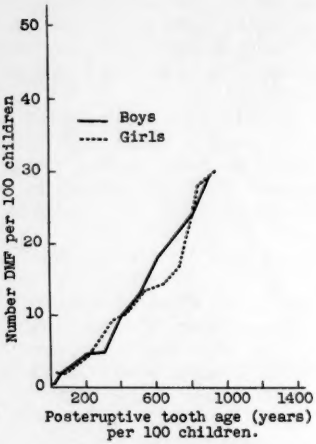


Figure 2
Maxillary 2nd premolars

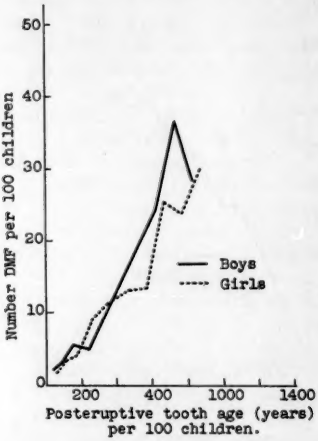


Figure 3
Maxillary central incisors

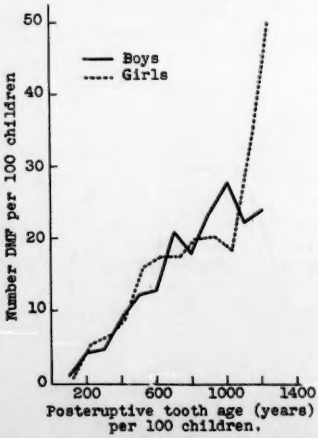


Figure 4
Maxillary lateral incisors

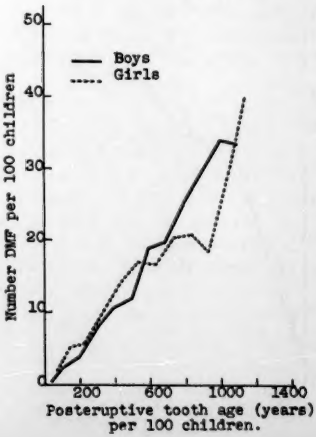


Figure 5
Maxillary 1st molars

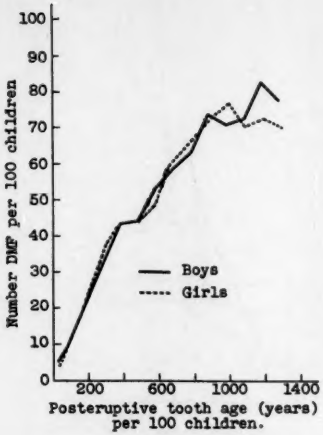


Figure 6
Maxillary 2nd molars

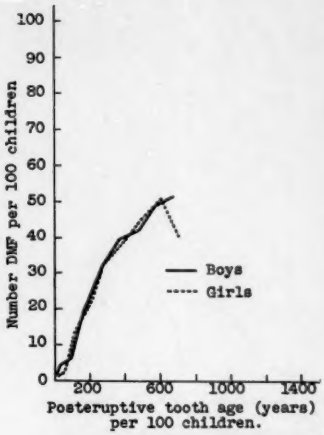


Figure 7
Mandibular 1st molars

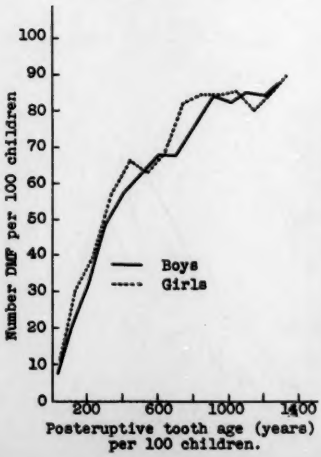


Figure 8
Mandibular 2nd molars

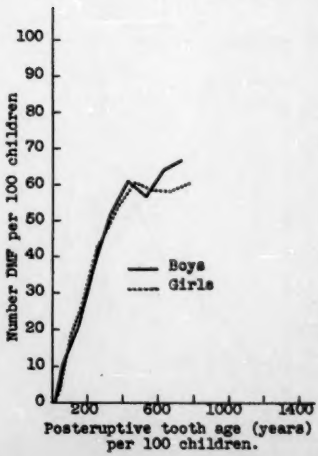


Figure 9
Mandibular canines

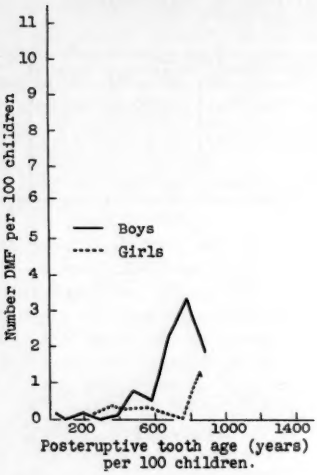


Figure 10
Maxillary canines

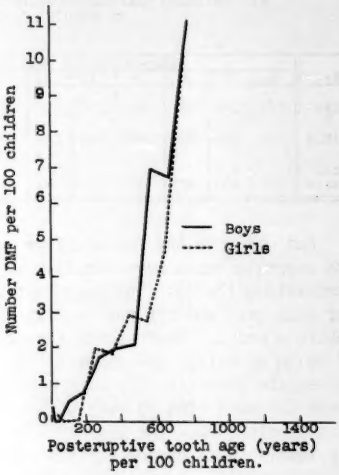


Figure 11
Mandibular central incisors

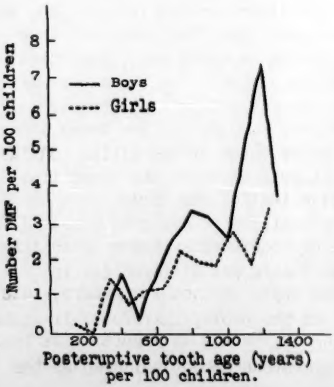


Figure 12
Mandibular lateral incisors

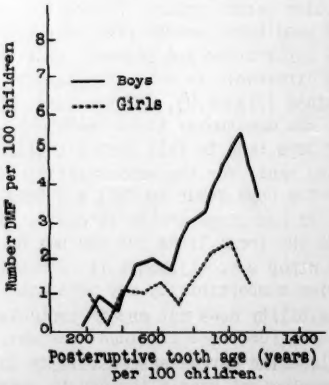


Table 1

NUMBER OF ELEMENTARY AND HIGH SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND AND ENVIRONS WHO RECEIVED DENTAL EXAMINATIONS. DATA ARRANGED BY SPECIFIED AGE AND SEX GROUPS

Sex	Chronological Age in Years (Last Birthday)														All Ages
	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Boys	171	197	231	253	270	262	299	272	341	325	193	172	74	27	3,087
Girls	156	206	256	240	259	269	297	293	354	326	252	183	74	5	3,170
Both sexes	327	403	487	493	529	531	596	565	695	651	445	355	148	32	6,257

Estimates of the posteruptive mouth ages of the individual teeth of the separate sexes were obtained by methods described in earlier communications (3, 4). The posteruptive tooth ages of the individual teeth for each sex, and for the several chronological age groups, are shown in Tables 4 and 5. These tooth ages represent averages for the single types of teeth of either the right or left sides of the mouth. The tooth age values are given in this form since the same type of tooth erupts at about the same time on each side. From the data shown in Tables 4 and 5 it is clear that the posteruptive ages of the single types of teeth are consistently higher for the girls.

In order to compare the caries susceptibility of each type of tooth in boys and girls, the tooth-specific DMF values given in Tables 2 and 3 were plotted against the corresponding tooth age values given in Tables 4 and 5. The results of this procedure are shown in Figures 1 through 14.

Taken altogether, the evidence presented in these several figures fails to reveal a consistent difference between the sexes in the caries susceptibility of the individual types of permanent teeth. For the maxillary first and second premolars, the maxillary central and lateral incisors, the maxillary first and second molars (Figures 1-6), the mandibular second molars (Figure 8), the maxillary canines (Figure 10), and the mandibular second premolars (Figure 14), the trend lines for boys and girls cross and recross, thus indicating no clear cut, consistent sex difference in susceptibility to caries attack. For the mandibular canines (Figure 9), and the mandibular central and lateral incisors, and the mandibular first premolars (Figures 11, 12, 13) the trend lines for boys tend to fall into a position above those of the girls. On the other hand, for the mandibular first molars (Figure 7) the trend line for the boys tends to fall a little below that of the girls.

If sex plays a role in caries susceptibility it would be expected that the trend lines for one sex would be consistently above those for the other sex. Although it is possible that a sex differential in caries susceptibility may exist for some teeth and not for others, this possibility does not appear probable. On the whole, therefore, it would appear from these findings that sex, per se, is not an important variable influencing caries susceptibility in Hagerstown school children of the chronological ages 6 through 19 years.

Figure 13
Mandibular 1st premolars

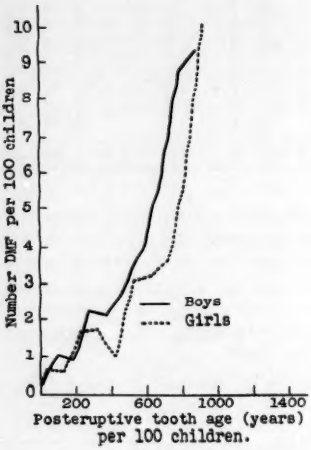
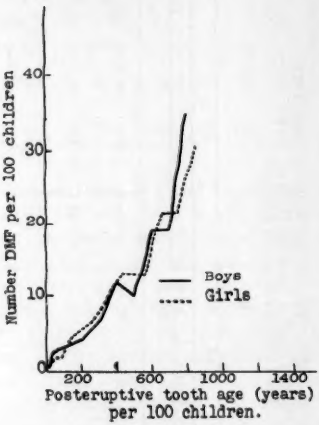


Figure 14
Mandibular 2nd premolars



REFERENCES

(1) Klein, Henry, and Palmer, C. E.: Studies on dental caries. VII. Sex differences in dental caries experience of elementary school children. Pub. Health Rep., 1938, 53, 1685-1690.

(2) Klein, Henry, and Palmer, C. E.: Dental caries in American Indian children. Pub. Health Bull. #239. Washington, U. S. Govt. Printing Office, 1937.

(3) Klein, Henry, Palmer, C. E., and Kramer, M: Studies on dental caries. II. The use of the normal probability curve for expressing the age distribution of eruption of the permanent teeth. Growth, 1937, 1, 385-394.

(4) Palmer, C. E., Klein, Henry, and Kramer, M.: Studies on dental caries. III. A method of determining posteruptive tooth age. Growth, 1938, 2, 149-159.

Table 2

NUMBER OF SPECIFIED TEETH AFFECTED BY CARIES EXPERIENCE (DMF), PER 100 CHILDREN OF SPECIFIED CHRONOLOGICAL AGES (LAST BIRTHDAY). DATA DERIVED FROM DENTAL EXAMINATION OF 6,257 ELEMENTARY AND HIGH SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, AND ENVIRONS.

Boys

Corresponding teeth of either side (right or left)	Chronological age (last birthday)													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Upper jaw													
Central incisors	--	--	0.87	4.15	4.63	8.78	12.21	12.87	20.68	17.85	23.58	27.91	22.30	24.08
Lateral incisors	--	--	0.44	2.37	3.71	8.02	10.54	11.95	18.62	19.69	26.13	29.36	33.79	33.34
Canines	--	--	--	0.40	--	--	0.50	0.74	1.76	2.00	2.08	6.98	6.76	11.11
1st premolars	--	0.26	0.44	0.60	2.22	3.25	4.68	4.97	10.12	13.08	18.14	21.22	24.33	29.63
2nd premolars	--	--	--	0.60	2.04	2.87	5.69	4.97	9.97	14.46	19.69	24.71	36.49	27.78
1st molars	4.68	9.65	20.57	31.82	43.71	44.09	53.68	59.01	63.49	72.93	70.99	72.97	83.11	77.78
2nd molars	--	--	--	--	--	1.53	4.52	6.62	20.53	32.31	40.16	42.44	49.33	51.85
	Lower jaw													
Central incisors	--	--	--	0.20	1.67	0.38	1.67	2.68	3.37	3.23	2.59	4.94	7.43	3.71
Lateral incisors	--	--	--	0.20	0.93	0.58	1.84	2.02	1.76	3.08	3.37	4.65	5.41	3.71
Canines	--	--	--	--	0.19	--	0.17	--	0.15	0.77	0.52	2.33	3.38	1.85
1st premolars	--	--	--	0.20	0.56	1.15	0.84	2.21	2.06	2.77	3.89	5.82	8.79	9.26
2nd premolars	--	--	0.44	0.40	2.22	2.67	3.35	3.86	7.04	12.00	9.85	18.90	18.92	35.19
1st molars	7.90	19.29	32.47	48.42	57.41	62.60	67.90	67.83	76.98	83.54	82.39	85.47	84.46	88.89
2nd molars	--	--	--	--	0.37	1.91	11.04	20.41	37.69	52.00	60.62	56.40	64.19	66.67

Table 3

NUMBER OF SPECIFIED TEETH AFFECTED BY CARIES EXPERIENCE (DMF), PER 100 CHILDREN OF SPECIFIED CHRONOLOGICAL AGES (LAST BIRTHDAY). DATA DERIVED FROM DENTAL EXAMINATION OF 6,257 ELEMENTARY AND HIGH SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, AND ENVIRONS.

Girls

Corresponding teeth of either side (right or left)	Chronological age (last birthday)													
	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Upper jaw													
Central incisors	--	--	0.59	5.42	6.76	8.92	16.16	17.24	17.23	19.94	20.04	18.31	32.43	50.00
Lateral incisors	0.32	--	0.39	5.21	5.60	10.04	13.98	16.73	16.53	20.25	20.84	18.31	28.38	40.00
Canines	--	--	--	0.21	0.20	--	--	2.05	1.84	2.92	2.78	4.65	10.14	--
1st premolars	--	--	0.20	2.09	2.32	3.53	5.56	9.22	10.45	13.50	14.29	16.94	27.71	30.00
2nd premolars	--	--	--	1.25	1.74	3.35	4.21	9.22	11.58	13.19	13.30	25.69	23.65	30.00
1st molars	3.53	12.14	22.66	37.29	44.21	44.43	49.50	61.44	66.39	72.70	76.79	70.77	72.98	70.00
2nd molars	--	--	--	0.21	0.20	1.49	3.03	14.00	22.32	33.74	40.28	45.90	50.68	40.00
	Lower jaw													
Central incisors	--	0.25	--	1.46	0.77	1.12	1.18	2.22	1.98	1.84	2.78	1.92	3.38	--
Lateral incisors	--	--	--	1.04	0.20	1.12	1.01	1.37	0.71	1.54	2.19	2.46	1.35	--
Canines	--	--	--	--	--	--	0.17	0.34	0.28	0.31	0.20	--	1.35	--
1st premolars	--	--	0.20	0.42	0.58	0.56	1.69	1.71	0.99	3.07	3.18	3.55	6.08	10.00
2nd premolars	0.96	0.25	0.39	1.04	2.90	3.16	4.38	6.66	9.18	13.35	13.10	21.59	21.62	30.00
1st molars	11.86	30.59	39.65	57.09	65.64	63.20	68.86	82.08	84.75	84.51	85.52	80.06	85.14	90.00
2nd molars	--	--	--	0.21	0.77	5.58	17.34	29.18	43.51	53.38	60.32	59.57	58.11	60.00

Table 4

NUMBER OF YEARS SPECIFIED TEETH HAVE BEEN IN MOUTH (POSTERUPTIVE TOOTH AGE), PER 100 CHILDREN AT SPECIFIED CHRONOLOGICAL AGES. DATA DERIVED FROM DENTAL EXAMINATIONS OF 6,257 ELEMENTARY AND HIGH SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, AND ENVIRONS.

Boys

Corresponding teeth of either side (right or left)	Chronological age													
	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5	17.5	18.5	19.5
	Upper jaw													
Central incisors	3.27	30.30	104.32	201.08	300.75	400.75	500.75	601.00	701.00	801.00	901.00	1001.00	1101.00	1201.00
Lateral incisors	1.12	8.82	38.10	101.22	190.06	288.35	388.31	488.00	588.00	688.00	788.00	888.00	988.00	1088.00
Canines	--	0.05	0.49	3.15	13.75	42.99	98.11	178.00	271.00	371.00	470.00	569.00	669.00	769.00
1st premolars	0.21	1.46	7.13	24.89	63.67	128.72	214.12	309.00	408.00	508.00	608.00	708.00	808.00	908.00
2nd premolars	0.13	0.85	3.90	14.00	38.71	84.19	153.27	238.00	334.00	432.00	533.00	632.00	732.00	832.00
1st molars	23.33	90.91	186.16	285.75	385.75	485.75	585.75	686.00	786.00	886.00	986.00	1086.00	1186.00	1286.00
2nd molars	--	--	0.01	0.12	1.45	9.46	37.73	98.00	184.00	280.00	380.00	480.00	580.00	680.00
	Lower jaw													
Central incisors	29.52	102.93	199.88	299.70	399.70	499.70	599.70	700.00	800.00	900.00	1000.00	1100.00	1200.00	1300.00
Lateral incisors	2.75	25.23	92.11	186.67	286.40	386.40	486.40	586.00	686.00	786.00	886.00	986.00	1086.00	1186.00
Canines	--	0.07	1.06	8.15	35.32	95.25	182.86	280.00	380.00	480.00	580.00	680.00	780.00	880.00
1st premolars	0.07	0.63	3.79	15.66	46.36	103.20	183.42	277.00	376.00	476.00	576.00	676.00	776.00	876.00
2nd premolars	0.14	0.78	3.50	11.98	32.54	71.91	132.98	213.00	306.00	403.00	502.00	602.00	702.00	802.00
1st molars	33.02	108.44	206.32	306.00	406.00	506.00	606.00	706.00	806.00	906.00	1006.00	1106.00	1206.00	1306.00
2nd molars	--	--	0.03	0.51	4.21	20.91	64.36	138.00	232.00	330.00	430.00	530.00	630.00	730.00

Table 5

NUMBER OF YEARS SPECIFIED TEETH HAVE BEEN IN MOUTH (POSTERUPTIVE TOOTH AGE), PER 100 CHILDREN AT SPECIFIED CHRONOLOGICAL AGES. DATA DERIVED FROM DENTAL EXAMINATIONS OF 6,257 ELEMENTARY AND HIGH SCHOOL CHILDREN OF HAGERSTOWN, MARYLAND, AND ENVIRONS.

Girls

Corresponding teeth of either side (right or left)	Chronological age													
	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5	17.5	18.5	19.5
	Upper jaw													
Central incisors	8.08	48.77	131.98	229.93	330.22	430.22	530.22	630.00	730.00	830.00	930.00	1030.00	1130.00	1230.00
Lateral incisors	1.12	12.08	55.03	137.77	235.07	334.64	434.64	535.00	635.00	735.00	835.00	935.00	1035.00	1135.00
Canines	0.01	0.10	1.05	6.84	27.98	75.15	152.83	246.00	346.00	445.00	545.00	645.00	745.00	845.00
1st premolars	0.06	0.77	5.55	24.89	72.20	148.94	242.65	342.00	442.00	542.00	642.00	742.00	842.00	942.00
2nd premolars	0.17	1.09	5.25	18.24	49.10	104.04	180.11	272.00	369.00	469.00	569.00	669.00	769.00	869.00
1st molars	25.88	98.64	196.05	296.10	396.10	496.10	596.10	696.00	796.00	896.00	996.00	1096.00	1196.00	1296.00
2nd molars	--	0.01	0.11	0.98	5.60	22.11	60.89	128.00	214.00	310.00	410.00	510.00	610.00	710.00
	Lower jaw													
Central incisors	58.40	142.01	240.32	339.68	439.68	539.68	639.68	740.00	840.00	940.00	1040.00	1140.00	1240.00	1340.00
Lateral incisors	6.54	42.43	121.60	219.27	319.20	419.20	519.20	619.00	719.00	819.00	919.00	1019.00	1119.00	1219.00
Canines	0.01	0.23	3.28	21.37	73.28	157.61	255.17	355.00	455.00	555.00	655.00	755.00	855.00	955.00
1st premolars	0.05	0.61	4.58	21.08	63.36	135.41	226.78	325.00	425.00	525.00	625.00	725.00	825.00	925.00
2nd premolars	0.18	1.04	4.65	15.86	42.18	90.16	159.12	246.00	346.00	446.00	546.00	646.00	746.00	846.00
1st molars	57.15	140.40	238.46	338.43	438.43	538.43	638.43	738.00	838.00	938.00	1038.00	1138.00	1238.00	1338.00
2nd molars	0.01	0.10	0.78	4.05	15.57	43.76	96.71	172.00	263.00	361.00	460.00	560.00	660.00	760.00